Quantification of interaction energies for host/guest peptides with a hydrated DMPC bilayer - A step towards membrane protein folding

Adams, Gareth

Award date:
1999

Awarding institution:
University of Bath

Quantification of Interaction Energies for Host/Guest Peptides with a Hydrated DMPC Bilayer - A Step Towards Membrane Protein Folding

Submitted by **GARETH ADAMS**

For the Degree of Ph. D.

of the University of Bath

1999

(Gareth Edward Adams)

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Quantification of Interaction Energies for Host-Guest
Peptides with a Hydrated DMPC Bilayer -
A Step Towards Membrane Protein Folding

Submitted by G.E. Adams for the degree of Ph.D.
of the University of Bath
1999

In this study the energetics of various amino acid residues in an all atom, fully hydrated DMPC bilayer are determined computationally and analysed. These energetics are a precursor to the generation of a solvation potential for membrane protein folding.

Molecular Dynamics simulations of hydrated pure DMPC bilayers demonstrate that a hexagonal initial geometry is better than a cubic geometry at reproducing experimental data. A poly-Ala peptide is used as a model of helical membrane proteins and inserted into the pre-equilibrated hexagonal bilayer. Mutations were introduced in the peptide at the bilayer interface and centre positions. A parallel and anti-parallel pure poly-Ala helix pair, plus an anti-parallel pair with the central residues mutated to Try and Tyr were also incorporated into the model membrane. After simulation, the structural analysis of the bilayer and the incorporated helices indicated that the energetic analysis of the system was valid. The bilayer was shown to adjust its structure to accommodate the incorporated helices in agreement with experiment.

The interaction energies give new insights into the interaction of charged, polar and apolar residues with the bilayer environment. The energetic analysis showed that the charged residues preferred the interface region, the aromatic residues preferred to be paired and the helix pairs preferred the anti-parallel arrangement, but Arg was observed to have unexpected stabilisation in the bilayer interior.

Acknowledgements

I would like to express my thanks to all the staff of Bath University, especially those of the Department of Chemistry, and to the following people for their help and assistance during my years at Bath University:

David Osguthorpe, my long suffering supervisor, for his patience, insight and advice on work and life in general. Without his support it is debatable whether I would have managed to get this far!

The past members of the M.G.U. who helped to develop my interest in chemistry, whilst making life at university interesting and always challenging - Andy Lemon, David Evans, Colette Maunder, Oz Parchment and Pnina Dauber-Osguthorpe.

Gerald Zapata-Torres for making the days shorter, managing to do what formal education failed to do, teach me a few words of foreign language, and for stimulating conversations. Also, for his help in developing the Fold Fit receptor modeling procedure.

The EPSRC for their financial support both for the Ph.D and for the visit to the University of Chile, Santiago.

Finally, I have to thank my wife Ali who, during the course of my studies, has been a source of constant strength and encouragement.

Contents

1. Introduction

Cells are the building blocks of life and the understanding of their function, structure and interactions are of fundamental interest to biomedical science. Human cells incorporate many different structures and molecules that define its existence. Central to its function are the membrane bound proteins - Integral Membrane Proteins (IMP) - which regulate transport of life-giving molecules and the expulsion or recycling of waste products. Without the presence of these proteins the membrane would only permit inter-cellular transport via permeation and other passive processes. Thus, understanding the basic frame work (the lipid bilayer) of the biological membrane and the interactions it has with the various solutes immersed in or adsorbed on its surface is a basic pre-requisite to understanding life.

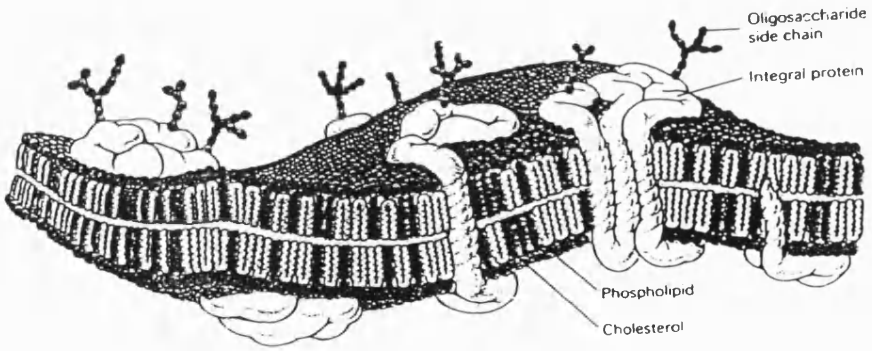


Fig. 1.1

Fig. 1.1 illustrates the complex nature of the cell membrane and shows that the phospholipid or lipid molecule is the main component of this dynamic milieu. The amphipathic nature of the lipid molecule gives the membrane its typical structure.

The problem of folding proteins has been known since the *in vitro* refolding of

haemoglobin in the 1930's, and the observation that it possessed the same activity as the native enzyme.[1] The unsolved IMP folding problem is beginning to attract a large amount of interest and numerous different routes are being taken in an attempt to better explain the phenomenon of how poly-peptide chains fold into well-ordered structures. Many questions still remain on the thermodynamic and kinetic basis for folding - do the amino acid sequences fold to a minimum in the free energy surface and through what mechanism do they achieve their folds?

The family of IMP's can be sub-divided into two main groups, according to their inclusion into the bilayer or membrane:[2]

(a) **Intrinsic** - These IMP molecules are observed to penetrate the bilayer and are thus bound to the lipid bilayer, but are able to diffuse along the bilayer plane as a result of the fluidity of the lipid molecules. A sub-set of this type of IMP, i.e. those that do not possess large non-membrane domains, are the focus of this study.

(b) **Extrinsic** - These IMP molecules are bound to the surface of the membrane via electrostatic and hydrogen bond interactions. This group will be discussed in this introduction but will not be studied further. This group contains the surface bound amphipathic helices.

The extrinsic IMP have a mixed hydrophilic and hydrophobic amino acid content, usually with the hydrophilic side chains on one face of the structure and the hydrophobic residues on the other. They gain a specific orientation to the bilayer via their electrostatic interactions with aqueous media and via their hydrophobic interactions with the upper region of the lipid bilayer.

As well as the central problem of protein folding in a membrane, other interesting associated research is being undertaken to better understand cell membranes. For

example, the fullerene family is being investigated as potential charge transporters across a membrane, as there is evidence that they can form photo-active aggregates.[3] The fluidity of erythrocyte membranes has been studied, with respect to the athletic ability of the human host and a clear correlation has been established between the lower membrane fluidity for sedentary individuals over sprinters and long distance runners.[4] The intuitive idea that capillary activity and micro circulation is fundamental to athletic prowess, has been confirmed.

The benefit of a better understanding of IMP structures and their interaction with the bilayer is being applied to many current medical problems and diseases. For example, extensive studies have been carried out into the basis for Alzheimer's disease via the structure of β-amyloid peptide and its interaction with the membrane. One theory, resulting from extensive studies into the membrane structure of this peptide, is that the membrane bound section undergoes an α-helical to β-sheet conformational shift which is the early stage of amyloid formation in diseased subjects.[5] It has also been shown that these peptides can form cation selective channels in planer lipid bilayers, which is predicted to cause cell death as a consequence of the resultant ion loss.[6] Thus, the development of drugs to prevent either channel formation or its activity, or to interrupt the α to β transition could perhaps help to reduce the effects of Alzheimer's and improve understanding of the neurotoxicity mechanism.[7]

Another practical example of such vital research is highlighted by studies on nimodipine, a drug that treats the neurological consequences of cerebral haemorrhaging. Studies of its partition coefficient into a brain membrane and its subsequent disassociation rate are helping to understand its potential efficacy in treating other neurological conditions, e.g. Alzheimer's.[8]

Amphipathic helices have been observed to lyse bacterial cell walls and not to harm the host cell. At critical concentration levels the peptides insert into the bilayer with a perpendicular orientation, which is proposed as the mechanism for cytotoxicity

and the selectivity of the lysis action is a consequence of the composition of the membrane.[9] The understanding of this process will have benefits in the development of antibacterial drugs. Whilst the lysis caused by the antibiotic peptides is a result of TM helices the lytic ability of melitin is related to its surface bound structure.[10]

Perez-Mendez [11] noticed that the class of peptides, known as the oblique orientated peptides, have a fusigenic activity when at a tilted position. Their example produced fusigenic action when at 40° but no activity was observed when orientated at 0°.

The life time of a typical IMP is complex beginning with mediated or spontaneous folding (discussed later) and then insertion into the host bilayer. The protein will normally then be active and functional. In some circumstances, there is evidence that they can be inserted into the bilayer in a dormant, inactive state which they remain in for a significant period of time and are then "activated" when there is a biological demand for them, i.e. the molecule is kinetically trapped.[12]

Thus, the field of lipid-bilayer protein chemistry and biology is vast and varied but the fundamental understanding of these complex environments is vital to both medical and pure science.

1.1. Integral Membrane Proteins?

Integral Membrane Proteins (IMP) are a very large group of protein structures that are associated with the cell membrane and can be divided into three main classes:[13]

1. The β-barrel structure (e.g. porin from Rhodobacter capsulate)[14]
2. The α-helical bundle structure (e.g. G-protein Coupled Receptors (GPCR))[15]
3. The single stranded helix (e.g. gramicidin).

The large number of structural studies on globular proteins has lead to the solution of many 3D structures, but relatively few structures of IMP's have been solved, and there are only a handful of membrane proteins families whose structures are in the

Protein Database (PDB), e.g. the porins.[14]

The two secondary structural types observed most, for membrane associated proteins that have had their structure solved, are the β-barrel and the α-helix bundle. The majority of the structures in the PDB Database are of the β-barrel type as numerous porin 3D structures have been reported whilst a lesser number of α-helical structures have been reported. In this study and many others the α-helical bundle is taken as the predominating secondary structure type of IMP's but, given the fact that the porins form the largest group of reported IMP structures, why should this be the case? These β-barrel porin structures must be very stable entities as they are found in the outer membrane of bacterial cells, which are very active systems and therefore they have to withstand potentially distorting interactions. This structural stability probably translates to their ability to retain their structure on removal from the bilayer, a process that is necessary for structural analysis. Conversely, the α-helical bundle arrangement is thought to be less robust as its stability is much more specific to the membrane environment and its structure is thus more easily lost on removal from the membrane.

Therefore, even though there are relatively few helical bundle proteins structures solved, conventional wisdom still says that the α-helical secondary structural type is predominant in IMP's as the membrane is observed to prefer structures that have fully internally satisfied backbone hydrogen bonding. Thus, in most IMP studies the α-helix is used as the basis for modelling or investigating membrane associated protein and peptide interactions. There are many families in the helical bundle IMP's. These are categorised by the type of fold they are observed or predicted to attain in the host membrane. The G-Protein Coupled Receptor (GPCR) family is one such group and they are very important proteins because they transduce signals across the cell membrane.[16] The GPCR's are designated to be transmembrane (TM) orientated proteins and to contain a bundle of 7 TM helices with extra-cellular loops and termini. The GPCR, 5-HT_{1A}, will form the basis of the semi-automated folding procedure in

Appendix I.

1.2. Integral Membrane Proteins - Classification Problem

Classically, the globular proteins have their structure classified to below 2 Å RMS via X-ray crystallography. This compares with the trans membrane (TM) section of bacteriorhodopsin (BRh), as studied by electron microscopy techniques, which has a horizontal resolution of 3.5 Å and vertical resolution of 10 Å.[17]

The use of X-ray methods to determine IMP structure is very difficult due to the fact that they tend to denature and lose significant amounts of native structure when removed from the bilayer environment. Using spectroscopic methods to study IMP structure gives distorted results due to the high degree of influence the bilayer exerts. Despite recent advances in technique such as NMR, IR, circular dichroism (CD) and neutron scattering detailed atomic level descriptions of these proteins is illusive. Thus, membrane protein science has focused on using semi-empirical, knowledge based and theoretical approaches to predict/model tertiary protein structures and to elucidate on their interactions. The computational techniques used in this study are amongst a group of modelling methods now widely used to either investigate protein structure or measure their interactions in these complex membrane systems at atomic level detail.

Trans-membrane helices are putatively α-helical but do contain a high number of residues that are commonly thought to be helix destabilisers or breakers (e.g. glycine) and β-sheet promoters (e.g. isoleucine, valine and threonine). It has been suggested that the inclusion of non-classical helix residues may occur in native membrane proteins as the residue could have a role in facilitating structural transitions in the membrane or in the vicinity of the membrane, allowing membrane insertion.[18]

The driving force for the process of protein folding is the interaction of the amino acid residues with its environment. Amino acids have a tendency to form one

secondary structure type in a particular dielectric medium and a completely different secondary structure in another medium of different polarity. For example residues 103-115 of ECORI endonuclease, ERE, when placed in Tri-flouro ethanol, T.F.E. (a helix promoter) are found to be mainly β-sheet in structure and when placed in sodium dodecyl sulphate, S.D.S. (a bilayer mimetic solvent) a highly helical structure was observed.[19]

This is also observed for the membrane peptides and proteins. Talli et al found that Saposin D in a membrane has an association dependent on the lipid and bilayer composition,[20] the peptide exhibited an ability to accommodate alternate secondary structure as a reaction to minor changes in its amino acid sequence (i.e. mutations) or solvent environment.[21]

A reasonable amount of information on general helical stability[22, 23] and the effects of the environment on the structural preference of the amino acid sequence exists. A huge amount of aqueous phase data is available but there is a dearth of detailed information on the properties of individual amino acids in TM structures as most studies are aimed at general and molecular level detail or properties.[24-27]

The fact that the membrane environment exerts a set of structural and interaction restrictions on the protein, compared to a globular protein in solution, enhances the scope for accurate structural analysis and prediction of intrinsic membrane proteins.

1.3. Inside-out Globular Proteins

The theory of globular protein folding is based upon the burial of hydrophobic residues into the interior of the protein structure where they are buried away from the solvating water molecules. Conversely, the polar or charged hydrophilic residues are exposed to the water molecules, satisfying their hydrogen bonding requirement. Even in this well studied area there is debate whether the process in driven by excluding hydrophobic residues or exposing hydrophilic residues. It is thought that the

hydrophobic exclusion is the driving force and the hydrophilic exposure is consequential - the hydrophobic effect therefore predominates.

The paradigm that most membrane proteins folding studies start with is the direct opposite of the globular protein scenario, that is the hydrophobic nature of the bilayer causes the polar or charged residues to bury themselves into the protein core and the hydrophobic residues to expose themselves to the lipid molecules.[28-30] This phenomenon, the inside-out protein, was investigated by Stevens and Arkin [31] who highlighted the fact that the 2 step model of protein folding conflicts with this postulated theory, as the first stage in the process requires individually stable helices. These isolated helices would contain polar or charged residues, which would be in contact with the lipid bilayer but the inside-out paradigm says that they would have to fold to a non-helical geometry. The analysis of the available database on IMP structure indicates that there is no correlation between the polarity of the amino acid and its position on the helix.[31] Even though the number of available structures is relatively small this finding does begin to contradict the theory as a comparative analysis of water soluble helix bundles reveal a clear correlation between hydrophobicity and solvent accessibility. Steven and Arkin concluded that helix bundling is better described by an optimisation of the Van der Waal (VdW) packing forces.[31]

This work has resonance with studies carried out to quantify the energy of interaction between the helices of BRh and the bilayer.[32] The resultant energies showed that even the polar residues had a favourable interaction with the solvating lipid bilayer, even a helix containing charged amino acids had favourable insertion energies. This suggests that the assumption inherent in the 'inside-out' paradigm, i.e. the bilayer is hydrophobic, is flawed and that the bilayer-peptide interaction is a complex one.

1.4. The Folding Pathway

Anfinsen in 1973[33] demonstrated that the program for globular protein folding

was contained in the primary amino acid sequence alone. Research into globular proteins has tended to confirm the theory[34] that proteins can fold to the native structure driven by the amino acid sequence alone. This situation becomes more complicated when the protein is membrane bound, as many factors contribute.

The process of folding is governed by many different criteria but the bilayer environment causes altered priorities for these criteria. The bilayer has a mainly hydrophobic character and the main driving force for secondary structure production is the need for the peptide to satisfy its hydrogen bonding network intramolecularly. The insertion of peptide into a bilayer is governed by a balance of the 'hydrophobic energy' and the loss of free energy caused by the desolvation of the peptide hydrogen bonds.[35]

The globular protein folds in a pseudo-random and non-hierarchical process as a complex amalgam of local and non-local interactions.[34] The IMP appear to undergo a semi-structured step-wise folding and insertion process - where the local interactions direct secondary structure folding as the protein sequence inserts into the membrane surface and then tertiary structure forms via long range interactions as the protein translocates across the bilayer.[36]

Stevens et al found that the process of TM peptide partition required their isolated helices to be formed either in-situ or at the surface of the membrane and to do the former the polar residues need to be soluble in the bilayer core.[37] Final aggregation occurs during TM orientation (the movement of the helices to a trans-membrane position) or after the TM location process. Thus, the tertiary structure must be a consequence of the Van der Waals (VdW) requirement of the residues. But localised electrostatic interactions can also be present to enhance the VdW interactions, for example the Light Harvesting Complex II protein has two helices held together by an ionic pair.[13]

Chipot et al used a water-hexane environment to study the folding of a undeca-leucine peptide and observed that it folds in the interface region of the solvent. The peptide then translocated into the hexane, suggesting that the IMP will form parallel to the bilayer and rotate to perpendicular orientation, with the N-terminus leading into the bilayer.[38]

As well as the step-wise methods the IMP fold can be formed via a translocase system where a peptide sequence is stitched into the bilayer via the N-terminus as it is synthesised[1]

A study on the residue distribution in 469 putative TM proteins found a maximum amino acid correlation at 3.6 residues per turn, characterising an α-helix as the predominant TM structure.[39] The α-helix is the classic IMP structure as it satisfies all its hydrogen bonding requirement internally, thereby minimising the backbone interactions with the bilayer. The β-barrel is the best approximation to an α-helix that a β-sheet can achieve as it wraps around the sheet edges, thus significantly reducing the exposure of the NH/CO atoms to the solvent.

Another important determinant in the folding pathway is the packing of the helices in the bundle structure. In the helix pairs often observed for small proteins or poly-peptides, the helices are usually parallel probably as a result of inter-helical hydrogen bonding or electrostatic interactions as it is accepted that the parallel arrangement causes unfavourable packing interactions. But in bundles with 3 or more helices the major arrangement is anti-parallel as this produces a good packing situation and aligns the helical dipole favourably.[40] The arrangement of helix pairs has two options, as a consequence of their tendency to 'coil', they can form either a L or R handed crossed packing. The L handed packing of TM helix pairs appears to be the more favoured of the two due to the more advantageous interlocking arrangement that results.[41]

Electrostatic interactions have also been identified as important in directing helix bundling and have been observed to hold the BRh structure together, with the hydrophilic residues directed inside the IMP and the hydrophobic residues toward the lipid molecules.[42] Factors such as VdW packing, anti-parallel arrangement and electrostatic interactions are thus important in stabilising helix bundles, highlighted by the fact that aggregation is observed to increase the degree of α-helicity of a bundled 4_6 peptide over the single helix 4_6 TM peptide.[43] Mitaku et al investigated the major factors causing or allowing the formation of a TM secondary structure and then to accomplish the tertiary structure. They denatured BRh using alcohol to the extent that only secondary structure remained and observed that the loss of tertiary structure highlighted the importance of hydrophilic residues in helix binding and bundle formation. The bundling was thought to be due to the exaggeration of such interactions in an apolar environment. Most helices are of a mixed nature and in which case their polar interactions are thought to take precedence.[44]

It can be concluded from the above studies, and the possible folding pathways described before, that when peptides approach a membrane they begin to interact with the surface, attracted by hydrophobic interactions. The formation of the TM helices that follows is driven by the quasi hydrophobic nature of the bilayer interface. The helices then orientate themselves via the N-terminal section into a TM structure with a concomitant formation of the helix bundle, the tertiary fold.

As well as the spontaneous methods, it has been shown that chaperone molecules can control the pathway to the native structure for both globular proteins and IMP.[45,46] Hunt et al studied the individual 7 TM helices of BRh in a phospholipid vesicle and found that 2 helices favoured non-native structure.[47] They concluded that the native structure folded spontaneously, as 5 helices reproduced their native structure when isolated in a bilayer environment, but intermolecular or interhelical or chaperone interactions are required for total folding. In general, chaperones are known to

mediate protein folding so as to prevent aggregation at intermediate phases and any subsequent mis-folding - they facilitate correct folding.[48]

1.5. Peptide - Membrane Interactions

Recently more effort has been directed at understanding the interaction of protein TM sections, model TM helices and actual bioactive peptides with a view to characterising the degree of helicity, using both experimental and simulation methods. These studies generally look at the influence individual residues have on the 'global' helix properties. An increasing amount of work is directed at the dynamics and properties of the individual amino acids, attempting to better explain the local environmental driving force and how this effects peptide or protein properties and folding. Studies are also being undertaken that use Molecular Dynamics (MD) to incorporate model peptides that represent or include amino acid segments that are observed experimentally or theorised to either partition, stabilise or de-stabilise helices, in model membranes.

Belohorcova et al [49] incorporated a 16-mer polyleucine peptide into a hydrated DMPC model bilayer, to investigate the hydrophobic matching of the acyl chain interior of the bilayer and the helix core. They also placed a lysine residue at the termini of the helix to 'anchor' the peptide to the interface of the bilayer. Using MD methodology they found that the central core (residues 4-15) were good helical structures with N and C-termini being very mobile and frayed. They observed a small tilt in the helix over the 1ns simulation, ≈16°, which maybe either a consequence of reduced bilayer equilibration or artefacts induced by the use of only 12 solvating lipid molecules. This small bilayer will enhance any Periodic Boundary Conditions edge effects and these may predominate in the trajectory.[49] They also used a relatively low cut-off for the non-bond interactions and a switching function, both are known to induce serious artefacts in the derivatives of MD trajectories.[50, 51]

Bondon et al investigated the structure of an amphipathic segment (residues 143-162) of the primary sequence of spiralin, using a combination of NMR analysis and MD. They found that the peptide was a random coil in aqueous phase but became helical on addition of a hydrophobic mimetic solvent, trifluorethanol. The helix was of good structure but with the terminal 3 residues exhibiting increased flexibility. Thus they found that designating a peptide as amphipathic is a good indicator of helicity when bound to the bilayer but they did not indicate the orientation of the peptide.[52] It has been observed that the amphipathic peptide melitin is surface bound,[53] thus these peptides are commonly seen to be surface orientated.

Hydrophobic residues (such as Ile, Val and Leu) are observed to be 60% of all the residues in the TM section of membrane proteins and peptides.[54] Russell et al observed experimentally that the helicity of a 25-mer host-guest peptide did not vary when the mutation guest was either Ala, Leu, Val or Ile and a 74±8% α-helicity was calculated for the group of peptides.[55] In an interesting polyalanine peptide simulation, Shen et al placed a 32-mer alanine blocked α-peptide into a DMPC bilayer that had been well equilibrated.[56] They observed a 'core' helical region, which exhibited good α-helicity between residue 12-24. This core stability degraded gradually towards the termini, with residues 8-12 and 25-28 being more mobile, i.e. less stable and the terminal regions being fully unfolded, existing as coils. They also found that whilst the core residues favoured the α-conformation, a degree of 3_{10} helical character was observed, which probably is a consequence of the bilayer preference for tighter molecules more in line with the lipid molecular structure.

Thus, from these peptide simulation and experimental studies it can be predicted that amphipathic helices tend to associate at the surface, with their hydrophilic/hydrophobic facial requirement satisfied through the biphasic nature of the interface. Hydrophobic helices orientate themselves TM to the bilayer and have good helical stability in the core of the bilayer, with frayed terminal residues in or near the

head groups caused by competitive hydrogen bonding.

1.6. Hydrophobicity Scales - Validity for Fold Prediction?

The mainly hydrophobic nature of IMP peptide segments has lead to the application of hydrophobic scales to help to attempt TM structure prediction in molecules that have only their primary sequence characterised. But the problems with such propensity scales is they do not take sufficient weighting from local environment and long range stabilising forces.[57]

Wimley and White published a membrane prediction scale for TM based on interfacial hydrophobicity, i.e. the preference of residue for the surface of a bilayer.[53] Their scale, for understanding the first step of peptide membrane inclusion, only takes into account the point preference of that residue and not the sequence dependent and partitioning effects of the heterogeneous bilayer. The scale suggests that most amino acids are indifferent to partitioning:

Tyr/Phe/Trp > Leu > Ile/Cys/Met/Gly/Val > Pro/Gln/Asn > His/Arg/Lys
Ser/Thr/Ala/His Asp/Glu

(Partitioning decreases to the right)

Chung et al studied a 27-mer polyalanine in a bilayer environment. They observed that the pure alanine peptide was weakly bound to the bilayer and precipitated into the buffer. When they mutated the 3rd/4th/5th residues to Leu the peptide bound to the bilayer in a TM orientation. Thus, small changes in the sequence of the peptide dramatically altered its behaviour, highlighting a weakness in general propensity scales. They proposed that the Leu may act as a helix nucleation site and therefore increase partitioning into the membrane but the increased hydrophobicity of the peptide on mutation to Leu is an additional factor.[58]

Another factor the general scales do not include is the effect peptide length has on the eventual structure. A leucine peptide has been observed experimentally to form

a good α-helix when its length corresponds to that of the bilayer interior but when it has a shorter length it prefers to form a β-strand.[59] This could be a result of the α-conformer 'burying' or reducing the exposure of the aliphatic side chains when the peptide is shorter. Conversely the longer length when in the β-conformer will expose side chains to the head group region.

The hydrophobic approach to prediction also does not take into account the effect charged/polar residues have on the general helix structure and on the residues distant from the charged/polar site. Gustafsson et al observed that the 21-mer (KL_4) peptide in a di-parmityl phosphatidyl choline (DPPC) and phosphatidyl glycerol (PG) bilayer achieved 84% helicity, surprisingly high for a peptide that includes such a percentage of charged residues. The peptide remained TM despite the accommodation of charged residues in the acyl chain region of the model membrane. It was postulated that the head group-Lys interaction compensates for the inclusion of the charged Lys residues in the apolar interior.[60] Such peptides are sometimes designated as an amphipathic helices and therefore should orientate parallel to the bilayer surface but due to its regular phasing of the Lys residues, which causes them not to be on the same face, it achieves TM orientation unlike regular amphipathic structures. A classic hydrophobic scale will disallow such peptide behaviour.

Another area of interest is the effect the aromatic amino acids have on properties of TM helices. They are highly hydrophobic according to the published scales but how do their bulky side chains interact with bilayers? Lee et al [61] found that an oligopeptide of Phe had a reduced affinity for the hydrocarbon phase of the bilayer therefore indicating that Phe or aromatic rings on the same helix are unable to match their hydrophobicity with that of the HC chains. The aromatic quadrupole, which results from the electron rich region above and below the ring, is a basis for favoured charge interactions at the bilayer surface[57] and thus explains the observed occurrence of these aromatic residues as helix or peptide anchors.

1.7. Quantifying the Energetics of Peptide - Bilayer Interactions

The energetic and thermodynamic basis of peptide-bilayer interactions have begun to be probed using both experimental and theoretically derived energies of interaction and partition for both helices and individual amino acids.

Table 1.1

$\Delta\Delta G_{part}$				
Val	Phe	Ala	Tyr	Thr
0.45	0.36	0.35	0.28	0.24

Deber and Li[54] were one of the first groups to look at the free energy of helix formation of a host guest alanine peptide, with Lys used to anchor it to the membrane. They reported the $\Delta\Delta G$ of partition relative to proline, see Table 1.1.

Interestingly, the Val had the highest stabilisation energy, especially because in globular proteins it is predicted to be of lower α content than Ala. But what is clear is that the membrane environment is less differentiating than the aqueous phase as all the values are closely related.[54]

Ben-tal[35] found that a 25-mer alanine peptide inserted into a bilayer and formed a helix, driven by the hydrophobic force, which was opposed by the entropic loss caused by trapping a helix in a bilayer. They calculated the free energy of solvation relative to water via electrostatic and non-polar components.

$$\Delta G_{solv} = \Delta G_{elec} + \Delta G_{np} = \approx -0.5 \text{ kcal mol}^{-1}$$

This is a shallow energy well therefore indicating subtle changes in the oligo-alanine peptide sequence could affect the equilibrium structure significantly. Their finding agrees with other structural and inclusion studies which shows that alanine is only a slight helix favouring amino acid.[35]

Russell et al used similar experimental methods to investigate the membrane affinity of aliphatic amino acids using a partition free energy term and related the affinity to position of the residue along the bilayer normal. They found that as mentioned earlier the host peptide achieved stable helicity for all guest residues but with the partition energies being different. Table 1.2 shows their values.

Table 1.2

Partition Energy, $\Delta\Delta G_{tr}$		
	Hgp	Ac
Ala	0.32 ±0.14	0.40 ±0.19
Val	0.99 ±0.17	1.12 ±0.17
Leu	1.56 ±0.17	1.77 ±0.12

Clearly the residues are of relatively similar partition energy, with the difference reflecting the hydrophobicity of the residues. Interestingly, the position of the residues in the helix relative to the bilayer head groups (Hgp) and acyl chains (Ac) has little affect on the values of the $\Delta\Delta G$, if at all they become a little more stable in the interior of the bilayer.[55]

This study begins to look at the position of the amino acid in the bilayer and therefore begins to add a little detail to the IMP folding problem. Interestingly even in the head group region the hydrophobic affect appears to predominate indicating that amphipathic surface bound peptides do not need to fully penetrate through the interface to stabilise their apolar face.

Woolf et al investigated the interaction energy of the 7 Bacteriorhodopsin helices and the three neutral versions of the charged helices, individually placed in a DMPC bilayer using MD.[62] They calculated interaction energies, i.e. solvation energies (ΔE_{solv}) and not the free energy of partition as they do not include the aqueous phase value. The analysis of the results indicated that there is a wide range of ΔE_{solv}, -5 to -30 kcal mol^{-1}, with the weaker interaction energies being dominated by the VdW

interactions and the stronger interactions by the electrostatics. The structural data indicate that the 'core' residues are stable as the RMS of the interaction energy is smaller than for the rest of the amino acids near the head groups.

Woolf et al[62] also observed that the charge of a particular amino acid affects the energetics of the residues other than just its own. This can be seen from the fact that the non-charged residues of the neutral and charged helices exhibit different ΔE_{solv} dependent on whether in the charged or un-charged peptide. Table 1.3 contains the energy averages for all the residues used.

Interestingly they observed that the Arg residue appears to have a similar value at a 0.5 relative depth in the bilayer to 0.35 relative depth. This result indicates that the Arg has a similar solvent environment at the Ac (acyl chain region) and HGp (head group region) positions in the bilayer (Table 1.3).

Table 1.3

ΔE_{solv} (kcal mol^{-1})		
	Hgp	Ac
Ala	-9 ±4	-4 ±1
Val	-9 ±5	-6 ±1
Phe	-18 ±5	-9 ±1
Tyr	-15 ±4	-10 ±3
Trp	-25 ±5	-16 ±2
Thr	-10 ±4	-6 ±2
Asp	-50 ±20	
Arg	-70 ±40	

Asp and Arg are low in population in the interior of the bilayer and have not been reported.

These values give another insight into the affinity of amino acids for the bilayer and the affects of adjacent charges on their energetics. Surprisingly, the Arg value appears to be anomalous as it exhibits similar affinity in two different positions.

1.8. Peptide-Bilayer Energetics and Thermodynamics

Wimley and White discussed the problems in accurately calculating thermodynamic data due to the difficulty in obtaining values for the folded or even unfolded peptide in aqueous phase, because IMP peptides are highly specific to the membrane environment. True bilayer thermodynamic data requires reversible denaturing of these highly stable entities, which often only denture at the exposed segments, e.g. loops.[53]

They postulate that 4 questions need to be answered in order to determine TM behaviour:

1. What structure do peptides adopt in a membrane environment?
2. What is the orientation they adopt and their location in the membrane?
3. What is the energetic cost of getting there?
4. What changes occur in obtaining the bilayer orientation?

The structural stability of IMP's is a balance of peptide bond partitioning into the bilayer and hydrogen bond formation. A 20-mer polyalanine peptide has a hydrophobic effect of ≈15 kcal mol^{-1},[53] with this stabilisation being counteracted by the cost in the formation of the hydrogen bonds. Two values of this cost have been reported, ≈ 0.6 kcal mol^{-1} per hydrogen bond[63] and ≈1.5 kcal mol^{-1} per hydrogen bond[35] thus giving 2 different ΔG_{partn} dependent on which is the correct hydrogen bond value. The correct value of this factor, in the bilayer, will be the determinant in TM partitioning. Due to the complex nature of bilayers it is important to study these parameters *in situ* rather than by using bulk solvent methods - bilayer are heterogeneous and complex solvents.

The hydrophobic nature of the TM helices has been studied in order to attempt to find a value of hydrophobicity that characterises the point at which a peptide will bind to a bilayer or not - the hydrophobic threshold to TM partitioning. It has been characterised at the poly-alanine helix level. [54] White et al point out that an accurate free

energy value for the TM partitioning is required for quantifying the 'threshold'.[64] A threshold to TM inclusion has also been calculated relative to the lengths of the hydrophobic core of the transmembrane peptide and the bilayer interior and it has been characterised that insertion will be favoured when the difference in the two lengths is less then 10 Å.[65] Killian et al probed this feature of TM peptides and described the basis on which the observed altered properties are based.[66] Thus, as this length difference can disallow TM orientation whilst allowing other orientations and because the translocation of proteins requires it to pass through the cell membrane, this phenomenon could be a route to a protein "sorting" mechanism.[66]

Therefore calculation methods are important as they allow the probing of the hydrophobicity level required for the inclusion of peptides into a membrane, the atomic level interactions of individual amino acids with the solvating hydrated lipid molecules and the individual contributors to the free energy of partition.

1.9. Prediction Methods

Whilst the characterisation of the structural and energetic stability of lipid bilayer - peptide/protein systems has been made, the characterisation of the fundamental interactions which constitute these energies is being probed to facilitate direct peptide-lipid modelling or IMP folding using mean field methods.

TM helix containing IMP's have been the focus of algorithms which predict the number and position of the TM sections but which are less accurate in predicting and disagree on the start and end residues. This, obviously is a problem for modelling studies using the predicted helices.[67] The problems with accuracy arise because the database, used as a basis for the algorithms, has relatively few known helical TM sequences in it.

The theoretical studies aimed at IMP structural prediction range from mean field methods[68] to single point mutation effect studies.[69] Lopez-Hernandez and Serrano

have developed an analytical equation incorporating solvent accessibility, hydrophilic area and dihedral angle terms which are fitted to experimental data. This analytical equation is used to predict the free energy consequence of a point mutation in the IMP structure, without knowing its tertiary structure. However this method, in not taking into account the protein tertiary structure, has no basis for including conformational changes which may result from the altered local environment of the IMP on mutation. Accurate predictions of the exposed surface is the determinant in their method and again the relatively low number of structures at the time of their study prevented further characterisation of the method.

Herzyk et al developed a rigorous automated method for predicting GPCR structure, using available experimental data. They employ a 4 stage modelling procedure, initially analysing the experimental and literature data to develop a protein representation and then they used restraints to penalise deviation from that representation. Finally they optimised the penalty function to produce a group of best fit structures. This method is extensive in its collation of and utilisation of literature data, but direct structural data on the 7 TM helix GPCR receptor is low. They originally used the BRh data as a basis for their work but in later work they used the more accurate electron density map of bovine rhodopsin to designate the TM structure and therefore develop a more representative model production method. But once again these methods are as good as could be achieved due to the low amount of 7 TM atomic level structure detail available.

Another restraint method being developed to build IMP molecules is IMPALA.[70] Its restraint field incorporates a distance description of the bilayer surface and atomic surface hydrophobicity dependent parameters. Thus, the restraint field describes the occlusion of hydrophilic residues from the lipid exposed face and inclusion of hydrophobic ones. MC simulation methods are used, assuming the bilayer core is made of CH_2 spheres. Unfortunately, the method is good for assessing known

structure but not as useful for predicting unknown structures.[70]

The mean field approach to protein structure prediction, using empirically derived solvent parameters, could hold the key to solving the IMP folding problem, allowing efficient prediction without predetermined and biased structural elements.

Self-consistent field theory (SCF) has been used to probe partition coefficients of aromatic and branched chain molecules in DMPC. The SCF theory reduces the many solvent molecules to a single molecule in an external field of mean force. An explicit set of equations, that describe the potential energy field of the central molecule, when solved at a fixed point are the SCF solution.[68] Other SCF examples use a pressure field to ensure that the bilayer system density remains near to that of liquid hexane.[71] Order dependent field theories have also been used, representing well the gel-liquid phase transition.[72]

The SCF method that Meijer et al[68] used was more complex using volume interactions, an anisotropic field, electrostatic and contact interactions, with no presumed structure except a lattice geometry. This method allows the development of physical system descriptors such as partition coefficients but as yet it does not allow atomic level structural detail as with MD. Thus, even though such SCF methods are CPU inexpensive they cannot give the required atomic level detail and insight into IMP structure required.

The use of knowledge based potentials to predict protein structure has also been evaluated by Rooman et al.[73] They used distance and dihedral potentials that describe the propensity of certain amino acid residue pairs to be a given distance apart or to populate a particular backbone dihedral angle range and was applied to probe IMP folds.[73] Again the main draw back in applying this method to IMP structural prediction is the lack of available fold data and thus knowledge based potential are highly context dependent - they will tend to be characteristic of the parent protein from which they are derived.

1.10. Working Hypothesis

Clearly the calculation of peptide or protein-bilayer interactions is a difficult problem to solve due to the lack of available good atomic level IMP structure, especially for those that contain helical secondary structure. Also the highly stable nature of the IMP tertiary structure in a membrane makes classical spectroscopic methods less able, than is the case for globular proteins, to directly solve the structure of these proteins.

Therefore, we propose to use MD calculations to model a di-mystirol phosphatidyl choline (DMPC) bilayer and then quantify the interaction energetics of specific amino acids in a α_R helical motif with that bilayer. This characterisation of the interaction energy of the amino acids could lead to future parameterisation of solvation potentials for the amino acids and thus, perhaps to the development of a simplified model for the interaction of the amino acids with a mean field bilayer model and allowing the studying of their secondary structural tendencies and their tertiary folding.

A polyalanine host peptide will be placed into a fully hydrated DMPC bilayer in a transmembrane orientation and mutated at three positions (2 at the bilayer surface and 1 in the interior) to include representatives of the main amino acid groups (described later) in order to study the interactions at different depths in the bilayer. A parallel polyalanine helix pair, an anti-parallel polyalanine helix pair and a anti-parallel helix pair including tryptophan and tyrosine residues will also be placed in the same bilayer. This will allow the studying of the effect of aligning like and unlike helical dipoles and the studying of the stabilisation afforded through the interaction of a pair of amino acids that can either hydrogen bond or interact aromatically.

Specific representatives of the amino acid groups are used due to computation time constraints as such large bilayer systems are very expensive.

2. Background Detail on Components of System

2.1. The Protein - Its Description and Structure

A protein is made up of a poly-peptide primary sequence covalently bonded (the peptide bond) amino acid units (residues) that forms local secondary structure elements such as a helix, beta sheet or random coiled structure which in turn coalesce into the final fold - the tertiary structure. The internal geometry of the amino acid units is not fixed and with the conformationally important dihedral torsion angles (ϕ, ψ, ω) being the defining parameter in the back-bone secondary and tertiary structures.

The $\phi\psi$ values tend to populate certain low energy regions in the $\phi\psi$ space, corresponding to the major secondary structure types:

$\approx -60°, -40°$ the α_R helix,

$\approx -120°, 140°$ β strand,

$\approx +60°, +40°$ the α_L helix (lesser populated).

$\approx -45°, -35°$ the 3_{10} helix (lesser populated).

The amino acid residues are able to form hydrogen bonds due to their polar groups which are relatively weak when compared to covalent bonds, but as they have directionality and specificity, they have an impact on protein structure. In a protein, the repeating secondary structures are present in the near optimal arrangement, in terms of ideal $\phi\psi$ values and in the best conformation for long range hydrogen bonding and steric fit.

These optimal values in the hydrogen bond arrangements have a significance in environments that have competing acceptor/donor species, requiring the secondary structure to optimise its hydrogen bonding network to compensate for the competing hydrogen bonds. Conversely, in the interior regions of the protein a small, almost negligible amount of competition exists.

2.1.1. Selected Amino Acids

The 20 α-amino acids commonly found in proteins can be grouped into categories dependent on either their C_β side-chain functional group or its functionality, below the groups are defined according to their side chain functional group (R) and are described below in three main categories:

1. Non-polar or Hydrophobic
2. Neutral polar
3. Charged Polar Groups

These main categories can be divided again into another group with mixed functional R groups, that have an aromatic nature and thus, the sub-set of aromatic side chains will be described separately. Each group is described below and each included amino acid is illustrated in Fig 2.1.

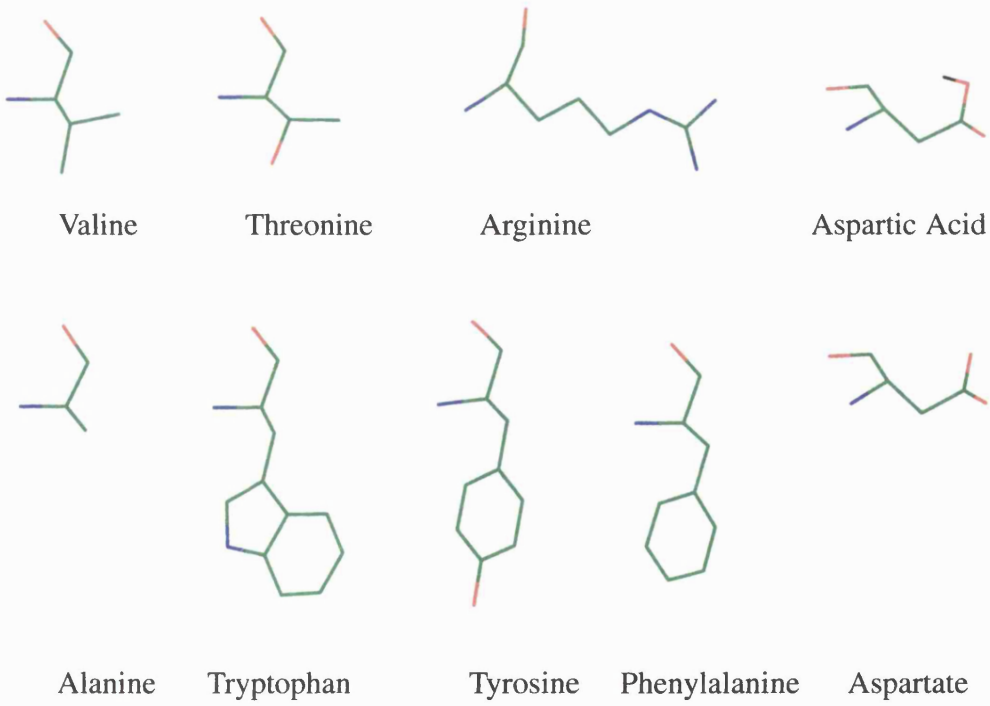


Fig. 2.1 Schematic of Amino Acid Structures used in study

2.1.1.1. Non-polar or Hydrophobic Group

The non-polar or hydrophobic group of amino acids contains Alanine, Valine, Leucine, Isoleucine, Proline, Phenylalanine, Tryptophan and Methionine. The non-polar group is represented in this study by the Alanine and Valine amino acids with Tryptophan and Phenylalanine being redefined into the aromatic sub-group. These residues are described briefly below.

Alanine - Ala (A)

The alanine amino acid is a very common and widely used as a control example, as it has a short hydrophobic side chain (see Fig. 2.1) with no chemical reactivity, no restricting conformational properties and is not specific to either structural face. Alanine has a relatively high propensity to form the α-helix globular proteins, as it satisfies all the requirements necessary to form a stable α-helix with a compact hydrogen bonding network and it also has been observed to be alpha helical favouring in membrane proteins. The alanine residue is, thus, the most prevalent amino acid in α-helices, as it has the smallest side-chain and requires no specific interaction to acquire the necessary $\phi\psi$ values, plus its backbone hydrogen bonding requirement is fulfilled.

Homogeneous, poly-alanine peptides form α helices in membrane environments, but these structures are only a shallow energy minimum in the energy surface. The incorporation of alanine, with its compact side chains, into such structures is probably to allow the main constituent amino acids, with their bulky side chains, to pack and attain the required structure, additionally Alanine adds to the hydrophobicity of the potential TM sequence without interfering with the secondary structure formation.

Valine - Val (V)

The valine amino acid with its hydrophobic and branched side chain is illustrated in Fig. 2.1, the bulky nature of the residue gives it a higher hydrophobicity than Alanine. Thus, the partitioning of the Val into the bilayer and globular protein interiors will be higher than Ala. Val favours β-sheet in globular proteins and slightly destabilises the helical geometry in TM helices of IMP due to the β-branch which confers conformational constraints in tight 3D structures, but they are common in general helical structures. The hydrophobic interaction of the side chain with the lipid bilayer may overcome the small destabilisation of the helix.

2.1.1.2. Neutral Polar Group

The Neutral Polar group contains Glycine, Serine, Threonine, Cysteine, Tyrosine, Asparagine and Glutamine. For the purpose of this study this group is represented by Threonine with Tyrosine being redefined into the aromatic sub-group with Phenylalanine and Tryptophan.

Threonine - Thr (T)

Fig. 2.1 shows the structure of the Thr amino acid which is described as a short side chain OH residue.

The hydroxyl group confers dual hydrogen bonding potential as it can hydrogen bond to either the backbone or to the solvent via donor or acceptor interactions. The OH is often observed, in central helix positions, to form a shared intra-molecular hydrogen bond with the i-3 NH and it is commonly allocated as a β favouring structure, because it favours the edge of β sheet allowing it to form solvent hydrogen bonds, the position it occupies next to the NH backbone can be assigned as helix destabilising. Thr is also observed to form a formal hydrogen bond with the first turn of an α helix, with the Thr acting as a N-capping residue. The methyl group of the side chain does confer a higher degree of hydrophobicity than the sister serine amino

acid and therefore it is observed to occur more frequently in the interior of globular proteins and confers higher lipid solubility.

Thus, it can be predicted from the structure of the Thr amino acid that it has a mixed hydrophobic and hydrophilic content and therefore the balance of these properties will determine the degree of bilayer interior and interface solubility.

2.1.1.3. Charged Polar Group

The charged polar group can be divided into to sub-groups according to their R group:

Positively charged - Basic Group

The basic group of amino acids contains Arginine, Lysine and Histidine, which are positively charged at pH 6.0, with the histidine also being a member of the aromatic subset. For the purpose of this study the group is represented by Arginine.

Arginine - Arg (R)

Fig. 2.1 shows the structure of the Arg with its long side chain and positively charged guanadinium tail.

Arg exhibits a good degree of order which is in contrast to the closely related Lysine amino acid which favours disorder. The Arg side chain can possess a hair-pin like structure, with a 'planar' hydrophobic surface facilitating a possible replacement for Isoleucine in protein structures.[74]

The guanadinium group at the end of the Arg side chain has also a planar network of 5 hydrogen bond donors which form near optimal water interactions but this arrangement is not totally favoured due to the concomitant entropy loss as a result of forming such a structure. This allows the possibility of hydrogen bond satisfaction via protein CO atoms. Arg does exhibit a preference for the last turn of a α helix (C-terminus) but does not often hydrogen bond to backbone CO atoms that are exposed at

the helix termini.

Thus, the Arg amino acid, as with the Threonine, has a mixed hydrophobic and hydrophilic structure and again it will be a balance of the long alkyl chain interacting with the core of the bilayer and the hydrogen bonding possibilities of the guanadinium with the interface, or other donor/acceptors, that will determine the solubility of the residue in the model membrane.

Fiori and Millhauser studied experimentally how a 3A4K peptide behaves as a consequence of Lys to Arg mutation. They found that the Arg mutant was more stable in the bilayer environment as it could hydrogen bond to amide CO atom at position i-3 position and therefore stabilise the first helix turn.[75] It is an interesting result indicating that the interaction of charged amino acids with peptide helices is not simple and as Arg are observed in protein helices will this mechanism stabilise TM peptides? They posed the question - how, as Arg residues are present in the active site of IMP's, can they stabilise their charge without ligand binding?

2.1.1.4. Negatively Charged - Acidic group

The acidic group of amino acids contains Aspartic Acid and Glutamic acid, which are negatively charged at pH 6.0 but neutral at pH 7. For the purpose of this study both the charged and the neutral version of Aspartic Acid are used.

Aspartic Acid - Asp (D)

Fig. 2.1 shows the structure of the Aspartic amino acid residue and clearly indicates the rotational symmetry of the terminal carboxylate terminal group.

The negatively charge confers an ability on the side chain to be a helix N-cap residue as it can form a hydrogen bond with the i+3 NH. Also this positioning at the first residue of a helix could be due to the favourable charge interaction with the macro-dipole of the helix,[76] which has a partial charge at the N-terminus.

The charged Asp residue is accommodated in globular proteins via either forming salt bridges in the interior of the protein or by forming solvent hydrogen bonds on the protein surface. In a membrane environment it is likely that the Asp charged example is strongly favoured at the interface but unfavoured in the hydrophobic interior of the bilayer. The neutral Asp residue will most probably exhibit a similar behaviour to the charged version but with a reduced magnitude, its functionality driven by the polar carboxylic acid group.

2.1.1.5. Aromatic Group

The aromatic sub-set of amino acids contains Phenylalanine, Tyrosine, Tryptophan and Histidine. For the purpose of this study Phenylalanine is used to represent the sub-set in the single helix simulation set of studies with Tryptophan and Tyrosine represented in the double helix simulation representative of an observed aromatic couplet in Bacteriorhodopsin.

Phenylalanine - Phe (F)

Fig. 2.1 shows the structure for Phe, Tryptophan (Trp) and Tyrosine (Tyr) residues and clearly demonstrates the aromatic nature of these amino acids but also the functional difference between them.

Phe is a completely hydrophobic residue with a strong preference for regular secondary structure. Conversely the Trp and Tyr side chains poses a single hydrogen bonding site which causes Tyr to be ambivalent about preferring the interior or exterior of the protein and the Trp to have a slightly reduced hydrophobicity.

These aromatic residues favour hydrophobic and protein-protein packing interactions with no significant preference for well stacked ring-ring structures but for more random clusters. An interesting cluster is the Tyr layer of the Photosynthetic reaction centre just below the outer membrane surface.[77]

These large rigid residues exhibit restrained motion in the protein structure, remaining close to their optimal $\chi_{1,2}$ values (for α helices, χ_1 = -60°, 180°). There is a third χ_1 value that is optimal structure for a mono-peptide but in a polypeptide is disallowed as it causes clashes with the backbone. The -60° energy well is the most favoured but is distorted to near -70°, with the 2 conformations being in fact equally populated.

Thus, the mixed hydrophobic and hydrophilic nature of the Trp and Tyr residues indicates that they again will have a bi-phasic interaction with the bilayer but will have a mainly hydrophobic interaction, which is the sole interaction for the Phe. The electron dense aromatic rings cause increased preference for the interface over that predicted for the other predominantly hydrophobic amino acids.

2.1.2. The Helix

The transmembrane polypetide sections of IMP's are either β-barrels or α-helices and almost all natural occurring helices are right handed (α_R). In α_R-helices the CO, NH defining interaction is through the *i,i*+4 residues which stabilise the typical $\phi\psi$ angles of -60°, -40°. Peptides can also form a 3_{10} helix with weaker hydrogen bonds between the *i,i*+3 residues, leading to a longer but thinner helix. The 3_{10} helix is less favoured and therefore is less common, except in α-helix C-terminal regions.[78, 79]

As well as the more classical α-helical and β-barrel structures the membrane environment also supports different geometries such as with Gramicidin A which has a single stranded, right handed motif, with a 4 Å pore and a 6.8 residues per turn.[80]

The use of techniques to investigate the folded state, its dynamic nature and structure, can be used to measure the inherent tendencies of an individual amino acid to effect the formation of secondary structure. This allows the estimation of the ability of an amino acid to initiate the formation of helices, σ and the ease of propagation of

an existing helix, S.[81,82] These parameters allow the investigation of the propensity of an alanine peptide to form an α-helix. The helix:coil transition in aqueous environments has been found to have only a small contribution from the formation of hydrogen bonds, as the favourable interactions of the polar atoms are canceled out by the loss of favourable solvent interactions and loss of entropy when locating the NH and CO entities (at room temperature $\Delta S = \approx 2$ kcal mol^{-1}).[82] The major driving force for the helix:coil transition in aqueous solution is a combination of the hydrophobic effect and the more favourable Van der Waal interactions of the α-helix, i.e. apolar effects.[82] But the unfavourable energy resulting from unsatisfied hydrogen bonding elements, in the interior of a protein or the bilayer, means that the formation of hydrogen bonds is still an important event in the folding of an α-helix. This is emphasised also from the fact that in a bilayer interior there are no competing solvent hydrogen bonding sites and no loss of favourable solvent interactions. Thus, the formation of the helix is favoured strongly in the bilayer interior as opposed to the aqueous solvent environment. Interestingly even though the helix is favoured in membrane environments the α-3_{10} transition barrier for a deca-peptide containing AIB residues in a membrane environment was found to have a ΔG=6 kcal mol^{-1}, which indicates that a change of helical geometry in membrane environments is facile.[83]

The coil-helix transition is also dependent on helix length, as it increases the number of long range stabilising electrostatic interactions increases.[79] This is confirmed by the folding of an extended poly-alanine peptide to an α helix which was found to proceed slowly when only VdW interactions were included, i.e. short range interactions and when long range electrostatic interactions were included the folding proceeded faster. Therefore, the formation of α helices is promoted by long range dipole interactions as well as the localised hydrogen bonding interactions.[84] This situation is also observed in bovine rhodopsin where the truncation of the polypeptide did not reduced the ability of the protein to form a TM structure. Therefore, the long

range helical interactions and local bilayer interactions are important in the structure of bovine rhodopsin.[85]

2.2. Phospholipid Bilayers and Model Membranes

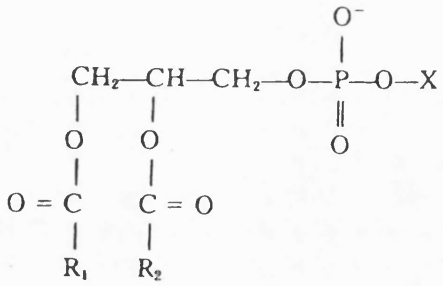
The phospholipid bilayer is the major component of cellular membranes and its formation is a direct consequence of the phospholipid's ability to self organise into vesicles when dispersed in water. The hydrophobic effect,[86] exerted by the water molecules, on the lipid, produces a lipid aggregate with the head groups in contact with the aqueous phase and the non-polar hydrocarbon chains directed inwards.

The structural, physical and dynamic properties of aggregated phospholipid assemblies have been widely studied using N.M.R.,[87-89] I.R. spectroscopy,[90,91] F.R.A.P.,[92] M.C.,[93] and MD,[94-96] with a view to understanding the dynamic and structural details that govern the complicated behaviour of membranes. As well as hydrated lipid bilayers, more complex, multi-component bilayers have been studied in order to investigate the effects, non-lipid molecules incorporated in the bilayer have on lipid structure and bilayer properties.[26,97]

2.2.1. Phospholipid and Membrane Structure

Phospholipids exist in two main classes, glycerophospholipids and sphingophospholipids, with the glycerophospholipids being the predominant phospholipid in the cell membrane. They have the general form of a head group linked to a phosphate group, esterified onto a glycerol backbone, with a fatty acid esterified onto the other two glycerol hydroxyls. The head group type and the hydrocarbon chain length attached to the glycerol characterises the phospholipid. Glycerophospholipids have three types of head group attached to the phosphate group, ethanolamine, serine and choline (Fig. 2.2).

Several crystal structures of phospholipids have been solved. The principal features of these structures are a "bent down" configuration of the head group and the non-equivalence of the hydrocarbon chains, with a kink being present at the Sn2, C_2 carbon position.[98,99]



Phosphatidylethanolamine —X = —$CH_2CH_2NH_3^+$

Phosphatidylserine —X = —$CH_2CH(NH_3^+)COO^-$

Phosphatidylcholine —X = —$CH_2CH_2N(CH_3)_3^+$

Fig. 2.2 General structural formulas for phospholipids, X - Head Group, $R_{1,2}$ - $Sn^{1,2}$

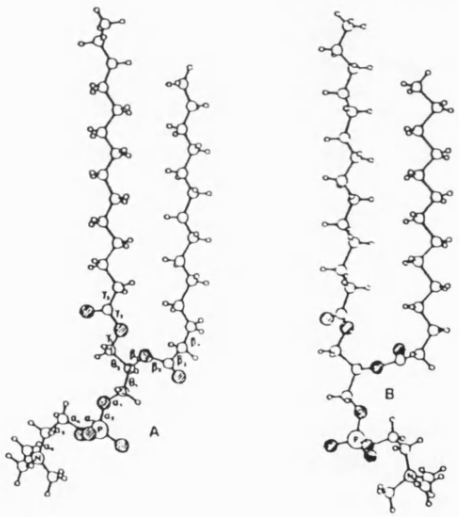


Fig. 2.3 Crystal structure of 1,2di-Mystirol-sn-glycero-3-phosphocholine.$2H_2O$.

Crystal structures of DMPC (Di-Mystirol-Phosphatidyl-Choline) are seen in Fig 2.3. The change to molecular configuration A from the B configuration allows the head groups to pack closer to the chain area/molecule. The main difference between the two molecules is in the polar head group orientation with respect to the hydrophilic

region of the molecules.

A dispersion of monomer lipid molecules will self assemble, aggregate and the aggregative properties of lipids are central in the formation of biological membranes.

(i) *(ii)* *(iii)*

Fig. 2.4 Aggregation states of aqueous phospholipids

(i) bilayer, *(ii)* micelle, *(iii)* inverted hexagonal cylinders

The unfavourable interactions of the hydrocarbon chains with water molecules, rather than the attractive tendencies of the adjacent hydrocarbon chains, are the driving force for the self assembly and aggregation.[86] This hydrophobic effect, or force, leads to the hydrocarbon chains coalescing, forming micelles and bilayers with the hydrophilic head group atoms exposed to the aqueous solvent.

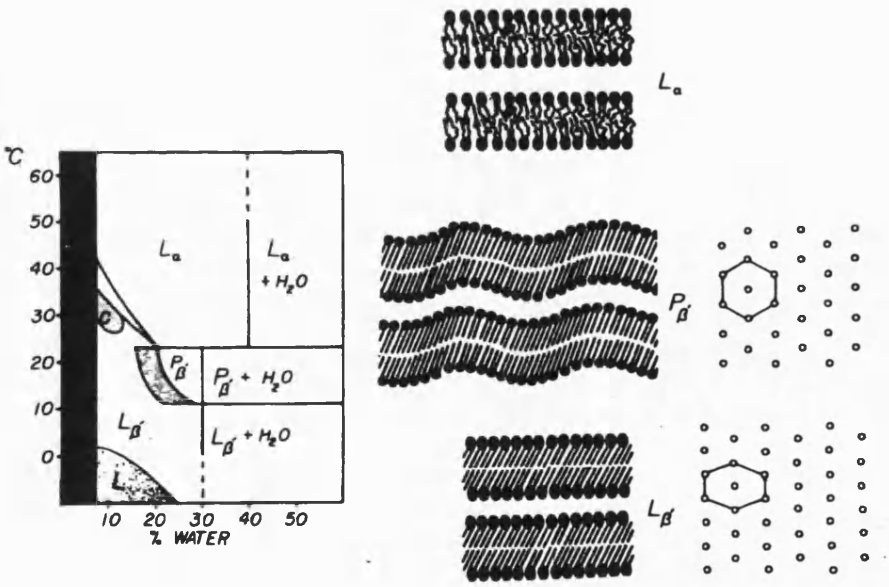


Fig. 2.5 Phase diagram for a hydrated DMPC bilayers.[100]

Homogeneous hydrated lipid bilayers have a well defined, thermotropic phase transition profile (Fig. 2.5), during which the hydrocarbon chains form a gel state, L_β, and a fluid liquid crystal state, L_α at higher temperature. As well as the two main phase stages, certain bilayers can also form a subgel phase, P_β' - a ripple phase (Fig. 2.5).

The transition temperature varies dependent on lipid type, but all the transitions are inversely related to the water content of the bilayer, i.e. as water content increases the phase transition temperature is reduced, up to a limit of around 25-30% water content.

At low temperature the solvated lipid bilayer exists in the crystalline L_c/L_c' state with similar molecular packing to anhydrous lipid crystals. At increased temperatures the hydrocarbon chains become more excited thermally and rotationally, overcoming the Van der Waal and steric interactions that promote crystallinity. When the rotational motion frequency increases the main transition occurs to the liquid crystal, L_α phase, the phase generally observed in bilayers at the physiological temperature.[101]

Tightly packed lipid bilayer structures are disrupted on hydration, causing a lowering in the bilayer density, as the inter-lipid head group hydrogen bonds are broken. A retention of the lipid molecular conformation is observed as the packing density is reduced, except the phosphate and choline groups, which are disrupted by hydration. Bilayers when hydrated are stable and unstressed.[102] The hydration of a lipid bilayer leads to a lower hydrocarbon chain density giving a corresponding enhancement in lipid dynamic motion. As solvation increases, the head group mobility and the frequency of CH_2 segmental rotational freedom is increased.

As a phospholipid bilayer consists of two distinct environments, the hydrophilic head group region and the hydrophobic hydrocarbon chain region, differential hydration is observed. The hydration of the chains is very low and hydration is restricted to the interactions at the interface. The much more important and more defining

interaction is the one between the head group and the aqueous solvent phase. Therefore, head group type is the key factor in the hydration process.

The water molecules close to the bilayer have significantly different properties when compared to bulk water due to perturbations, created by the interface, in the highly ordered hydrogen bonding structure of the water. The head group has a direct effect on water molecules in the adjacent layer.[96] This results in an indirect perturbation of the outer layers, as they rearrange to accommodate the ordering effect of the head groups on the inner water layer.

The rotational isomerisation of the C-C bond in the di-acyl hydrocarbon chains, gives the chains their characteristic motion and therefore, their dynamic properties. This dynamic motion is rapid and is transferred into the rotational and translational lipid mobility that categorises membrane fluidity.

The CH_2 segmental mobility is a balance of entropy, S and enthalpy, H. The rotational isomerisation of a C-C bond has 3 allowed conformations, gauche$^{\pm}$ ($\pm$120°) and trans (180°), with each having a biasing effect on the balance of entropy and enthalpy. The gauche isomer, the kink forming isomer, has an unfavourable effect enthalpically, but is entropically favourable, as it introduces conformational flexibility. This balance of entropy and enthalpy produces a uniform population of rotational isomers along the chain, because the increasing enthalpy penalty for the formation of a gauche isomer, is compensated for by the increasing disorder a gauche isomer creates.

The statistical probability of finding gauche bonds in the hydrocarbon region equals 4 per chain but, the experimental average varies from 3.5[103] to 2.6[104] gauche bonds per chain. The cooperative nature of chain rotation gives rise to the characteristic fluid property of lipid bilayers and equilibrium population of gauche rotamers, which in turn govern the structural and thermodynamic properties of a phospholipid bilayer.

The head group polar interactions of aggregated phospholipid molecules contribute directly to bilayer or membrane function, both the short range interaction of the adjacent lipid head groups, and the long range interactions of head groups across the bilayer or with head groups across the inter-bilayer water, in multi-lamellar aggregations. The head group, whilst being flexible, has preferences for particular torsional combinations and certain torsions, e.g. $\alpha_{1\text{-}4}$ are related to the crystal structure dihedrals. The tendency for them to be in the gauche state, relates to the inherent minimising of the coulombic interactions of the adjacent head groups.[104]

The head group dynamics can be categorised by the use of the angle vector joining the phosphate and the nitrogen, making an angle with the bilayer normal. The P-N vector angle is aligned with the bilayer surface but directed into the solvent at an angle of $\approx 20°$.[105] As well as spectroscopic analysis techniques, MD has also been used to investigate the head group interactions with water, by the calculation of the radial distribution function of water molecules from specific head group atoms or entities.[96, 101, 106]

Damodaran et al[96] used the MD simulation of a DMPC bilayer to investigate the type of water structure around the head group. They found the $N^+(CH_3)_3$ entity in the DMPC bilayer, to be hydrophobic in nature, causing the waters to form a clathrate type structure around the head group, even though the choline group has a formal positive charge.

The bi-phasic structure of the bilayer, resulting from the fluid acyl chain region and the more crystalline head group region, produces an interface region as the hydrophilic character of the head group region merges into the hydrophobic bilayer interior region. This interfacial region is diffuse, possessing no easily definable plane of action, but has a densely packed nature, with consequences for trans bilayer diffusion. The overcoming of the interfacial interactions and the high density of the region, is the rate limiting step for the diffusion of charged and uncharged species.[22]

There are 3 forces that make up the interfacial effect - the hydrophobic contribution, which is determined by the surface area of the hydrocarbon region that is exposed to water, the repulsive contribution arising from the electrostatic interactions of the polar head groups, and the steric contribution between, both the adjacent polar head groups and the phospholipid chains (Fig 2.6). In a bilayer, the resultant of these forces is at an equilibrium and gives rise to the interfacial free energy. The forces that produce the interface, the two repulsive components and the attractive component, act through points displaced from the hypothetical interface giving the bulk bilayer a propensity for curvature.[107]

The calculation of electrostatic profiles[94, 106] gives a measure of the limits of the interface and an approximate definition of the interface is taken as the peak in the electron distribution. But the diffuse nature of the interface in reality results in a definition determined by the effect it has on the nature of water. Thus, the use of the water density profiles[94] gives an indirect measure of the interfacial length. A definition of the boundaries of the interface is taken as the plane through the points where the water density drops, from 90% to 10% and is found to be ≈ 10 Å in width.[94]

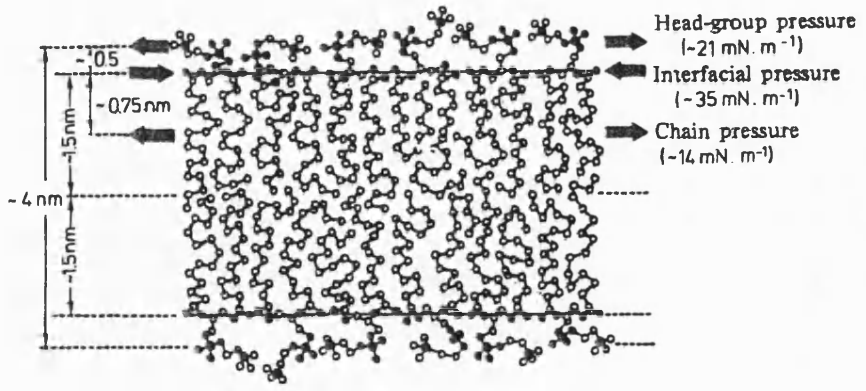


Fig. 2.6 Schematic representation of the forces in lipid bilayer.

2.2.2. Bilayer Models

A large amount of work has been directed towards the refinement of bilayer models and the characterisation of various bilayer structural and dynamic properties, e.g. the fluid hydrocarbon chain interior, the quasi-crystalline head group and the degree of water penetration onto the interfacial region. MD simulations have been used to attain working bilayer models which will reproduce experimental detail whilst allowing the investigation of properties that are difficult or impossible to obtain experimentally.

Recently several papers have been published that use classical MD to study the ability of methodologies and force fields to reproduce bilayer, and therefore, membrane environments. Gabdoulline et al studied a DMPC bilayer of 36 lipid molecules. They observed a surface area (SA) per lipid head group of 50$\AA^2$ which was significantly below the experimental value of 65-69$\AA^2$,[108] they also observed a 0.22 gauche fraction at the lower end of the experimental range, 0.2-0.3.[109] These lower bilayer parameters could be explained by the relatively small bilayer which will experience residual edge effects from the periodic boundary conditions and the low level of head group hydration.

The major problem with accurate reproduction of biological membranes by model bilayer systems is their mulit-component profile and their capability for structural heterogeneity. Lipids in membranes also have a tendency to occur in long lived sub-domains of gel and L_α phases. But at present incorporating its many minor parts into the model membrane are not feasible due to the need for the simulation to be time efficient and the problem with obtaining stable multi-component systems. As many dynamic functions of bilayers are characterised by relaxation times of 100 ps to 1 hr, very long simulations are required in order to equilibrate such multi-component bilayers.[110]

At the moment, the validity, or not, of using a surface tension (ST) factor in the

pressure term of MD simulated bilayers is a matter of debate in the current literature. Some studies published have used a pressure calculation adjusted to accommodate ST whilst others have postulated that, as ST is defined as a derivative of free energy with respect to area (per molecule) and in a bulk, equilibrium membrane the lipids will adjust their area so that the ΔG is at a minimum, thus ST has no value.[111, 112]

Chiu et al studied DMPC bilayers using a MD methodology that used a value for the lateral pressure equivalent to ST of a monolayer in a Langmuir Trough.[113] Their use of a monolayer derived lateral pressure gives reasonable but lower gauche angle population and surface area (SA) values. This is reflected in the fact that the ST used was -100 atmospheres and even though it achieved a stable model it did not reproduce the physical results as accurately as the non-ST simulation above which gave a SA of 61.8$Å^2$ as opposed to 57$Å^2$ (literature value 62$Å^2$).[114] Tieleman et al also concluded that the use of a ST term does not produce different values to constant pressure for the characteristic bilayer properties. [115]

Robinson et al[104] analysed the MD trajectory of a DMPC hydrated bilayer with a view to understanding head group structural properties. Their simulation was able to reproduce some experimental details, e.g. the ^2H order parameters of the acyl chains, the number of gauche bonds per molecule and certain head group torsion values. It was unable to reproduce the required degree of head group rigidity that experiment suggests. The starting structure was thought to be of fundamental importance to the resultant structure, as well as the parameterisation and the long range coulombic interactions. The starting structure should include a head group orientation related to those of the crystalline state and the acyl chain conformation related to the L_α phase structure.

As well as the explicit atom MD simulation model other reduced representation and lattice model based methods have been used to reproduce experimental data. Whilst they allow the calculation of longer time scale data, MD simulations, due to their inherent complexity, provide a more faithful model of hydrated lipid systems. However, they are restricted in the time scale that can be sampled, thus, they are less able to reach the large relaxation times of some dynamic and structural properties, such as diffusion. Milik et al[93] developed a tetrahedral lattice model that uses M.C. dynamics to give a simulation on the nano-second time scale. The model has good qualitative agreement with available data. Mean field methods are also used to allow the simulation of longer, micro-second time scales.[116]

In mean field studies only a single molecule or part molecule is maintained explicitly and a potential of mean force, along with appropriate random forces, is used to reproduce intermolecular interactions of the explicit molecule. The global environment is described by field parameters, which are adjusted to reproduce the order parameter of the methylene groups and the frequency dependent ^{13}C N.M.R. relaxation time, τ_1.

The diffusion of water through a membrane is a multi-stage process, with permeability rates in the bilayer being position dependent and thus cannot be thought of in terms of a homogeneous solubility. The diffusion rate of water molecules in a bilayer is at a minimum in the region of highest lipid density and highest in the central region, where the acyl chain density is low, thus, the rate limiting step is the movement through the dense region. These findings lead to the development of a four-region model scheme[95] to over-come the inadequacies of the two-phase liquid alkane/water model.

Region 1 : Low Head Group Density - extends from the membrane perturbed waters, to when the water density approximately equals that of the head groups. This region is likely to be very diffuse as the head groups have long range effects on the bulk water.

Region 2 : High Head Group Density - the internal boundary is at the point that the water density drops to below 1% with no existence of bulk waters, ≈ 7.5 Å in width.

Region 3 : High Tail Density - this region begins at the end of the penetrating water strands, with the inner boundary being at the point that the lipid density is equal to that of liquid hexadecane. The bilayer density in high through this region, 7 Å in width.

Region 4 : Low Tail Density - is this region the lipid density is low, spanning both of the monolayers, and equals that of hexane, with a width of 11 Å.

Defining the interfacial region as an area combining regions one and two gives an interfacial width of 11 Å and therefore, 40% of the membrane is involved in the interfacial interactions.

One problem with some MD models used in the evaluation of physical properties such as the diffusion rate, is the use of a united atom carbon potential[26, 104] combining both the carbon and hydrogen potentials, with an increased Van der Waals radius used. This lack of explicit hydrogens reduces the friction that is exerted on a passing molecule, leading to a enhanced diffusion rate. Therefore, for increased accuracy of the diffusion calculations all hydrogens must be explicitly included.[22]

2.3. The Hydrophobic Mismatch

A significant amount of work has been carried out to investigate the so called - "hydrophobic mismatch" phenomenon, i.e. what is the consequence of having a disparity in the hydrophobic length of the peptide and the host bilayer? de Planque et al studied the affect of incorporating a WALP peptide into bilayers with different length acyl chains (DLPC, DMPC, DPPC, DSPC). They observed that the bilayer thickness was related to the degree of mismatch and caused a concomitant change in the lipid head group area.[117] They concluded that the change in bilayer thickness was due to an adaptation in the acyl chain order but this adaptation was insufficient to relieve all the mismatch, but there is evidence that the bilayer can tolerate a degree of mismatch.[56, 59]

Hydrophobic mismatch can cause numerous changes in the relationship between peptide and bilayer and can affect the activity of IMP's indicated by the fact that altering the host bilayer thickness can cause a change in the activity of the (Na+,K+)-ATPase.[118]

The mismatch therefore results in local domains being developed which act to minimise unfavourable local interactions. The incorporation of peptides into a bilayer has a direct effect on the ordering of the adjacent lipid molecules, and when the hydrophobic length is less than the lipids they reduce the order of the lipid acyl chains and when they are longer they increase the lipid acyl chain order.[119]

References

1. G. V. Heijne, *Bioessays*, **71**, 25-30 (1995).
2. S. J. Singer, *Annual Review Biochemistry*, **43**, 805 (1975).
3. S. Nui and D. Mauzerall, *Journal American Chemical Society*, **11**, 5791-583 (1996).
4. K. Thornton, *Journal Biological Chemistry*, **268**, 19906-14 (1993).
5. J. Talafous, K. J. Marcinowski, G. Klopmann, and M. G. Zagorski, *Biochemistry*, **33**, 7788-96 (1994).
6. N. Arispe, E. Rojas, and H. B. Pollard, *Proceeding National American Society*, **90**, 567-571 (1993).
7. T. Kohno, K. Kobayashi, T. Maeda, K. Sato, and A. Takashima, *Biochemistry*, **35**, 16094-16104 (1996).
8. L. G. Herbette, P. E. Mason, K. R. Sweeney, M. W. Trumbore, and R/. P. Mason, *Neuropharmacology*, **33**, (2), 241-249 (1994).
9. H. W. Huang, *Journal Physics II France*, **5**, 1427-1431 (1995).
10. M. Monette, M. R. Van Calsteren, and M. Lafleur, *Biochimica et Biophysica Acta*, **1149**, 319-328 (1993).
11. O. Perez-Mendez, B. Vanloo, A. Decout, M. Goethals, F. Peelman, J. Vandekerckhove, R. Brasseur, and M. Rosseneu, *European Journal of Biochemistry*, **256**, 570-579 (1998).
12. S. Arumugan, S. Pascal, C. L. North, and et al, *Proc. Natl. Acad. Sci. USA*, **93**, 5872-5876 (1996).
13. W. Kuhlbrandt and R. B. Gennis, *Current Opinion in Structural Biology*, **4**, 517-518 (1994).
14. A. Kruesch and G. E. Sulz, *J. Mol. Biol.*, **243**, 891-905 (1994).

15. G.F.X. Schertler, C. Villa, and R. Henderson, *Nature*, **362**, 770-772 (1993).

16. S. Watson and Arkinstall, *The G-protein linked receptor In Factsbook,* Academic Press, London (1994).

17. R. Henderson, J.M. Baldwin, T.A. Ceska, F. Zemlin, E. Beckmann, and K.H. Downing, *J. Mol. Biol.*, **213**, 899-929 (1990).

18. S. C. Li and C. M. Deber, *FEBS Letters*, **311**, 217-220 (1992).

19. L. Zhong and W. C. Johnson, *Proceedings of the National Academy of Sciences of the United States of America*, **89**, (10), 4462-4465 (1992).

20. M. Tatti, R. Salvioli, F. Ciaffoni, P. Pucci, A. Andolfo, A. Amoresano, and A.M. Vaccaro, *European Journal of Biochemistry*, **263**, 486-494 (1999).

21. L.P. Li and C.M. Deber, *Biopolymers*, **47**, 41-62 (1998).

22. T. R. Stouch, D. Bassilino, and H. E. Alper, *Cray Channels*, **16**, (1), 12-17 (1994).

23. M. L. Smythe, S. E. Huston, and G. R. Marshall, *Journal of the American Chemical Society*, **115**, 11594-11595 (1993).

24. J .M. Steim, O. J. Edner, and F. G. Bargoot, *Science*, **162**, (909) (1968).

25. H. L. Scott and S. Kalaskar, *Biochemistry*, **28**, 3687-3691 (1989).

26. A. J. Robinson, W. G. Richards, P. J. Thomas, and M. M. Hann, *Biophysical Journal*, **68**, 164-170 (1995).

27. Y. P. Zhang, R. N. A. H. Lewis, G. D. Henry, B. D. Sykes, R. S. Hodges, and R. N. McElhaney, *Biochemistry*, **34**, (7), 2348-2361 (1995).

28. Eisenberg and R. Wall, *Journal Molecular Biology*, **179**, (1), 125-142 (1984).

29. D. C. Rees, L. antonio, L. Atonio, and O. Eisenberg, *Science*, **245**, (4917), 510-513' (1989).

30. D. M. Engleman and G. Zaecai, *Proceedings National Academy of Science*, **77**,

(10), 5894-98 (1980).

31. T. J. Stevens and I. T. Arkins, *Proteins Structure Function Genetics*, **36**, 135-143 (1999).

32. T. B. Wolf, *Biophysical Journal*, **74**, 115-131 (1998).

33. C. B. Anfinson, *Science*, **181**, 223 (1973).

34. Y. W. Chen and A. R. Fersht, *FEBS Letters*, **347**, 304-309 (1994).

35. N. Ben-Tal, *Biophysical Journal*, **70**, 1803-12 (1996).

36. J. L. Popot, *Current Opinion Structural Biology*, (3), 532-540 (1993).

37. T.J. Stevens and I.T. Arkin, *Proteins-structure Function And Genetics*, **36**, 135-143 (1999).

38. C. Chipot and A. Pohorille, *Journal of The American Chemical Society*, **120**, 11912-11924 (1998).

39. F. A. Samatey, C. Xu, and J. L. Popot, *Proc. Nat. Acad. Sci. USA*, 4577-4581 (1995).

40. W.D. Kohn and R.S. Hodges, *Trends in Biotechnology*, **16**, 379-389 (1998).

41. D. Langosch and J. Heringa, *Proteins-structure Function And Genetics*, **31**, 150-159 (1998).

42. P.E. Volynsky, D.E. Nolde, A.S. Arseniev, and R.G. Efremov, *Internet Journal of Chemistry*, **2**, 1-17 (1999).

43. S. Sakamoto, H. Mihara, and E. Matsuo, *Bulletin of Chemical Society of Japan*, **68**, 2931-39 (1995).

44. S. Mitaku, K. Suzuki, S. Odashima, and H. Itoh, *Protein: SF&G*, **22**, 350-362 (1995).

45. R. J. Ellis, *Current Biology*, **4**, (7), 633-635 (1994).

46. C. A. Kumamoto, *Molecular Microbiology*, **5**, 19-23 (1992).

47. J.F. Hunt, T.N. Earnest, O. Bousche, K. Kalghatgi, K. Reilly, C. Horvath, K.J. Rothschild, and D.M. Engelman, *Biochemistry*, **36**, 15156-15176 (1997).

48. A.L. Fink, *Physiological Reviews*, **79**, 425-449 (1999).

49. K. Belohorcova, J.H. Davis, T.B. Woolf, and B. Roux, *Biophysical Journal*, **73**, 3039-3055 (1997).

50. H. Schreiber and O. Steinhauser, *Biochemistry*, **31**, 5856-5860 (1992).

51. H. Schreiber and O. Steinhauser, *Chemical Physics*, **168**, 75-89 (1992).

52. A. Bondon, P. Berthault, I. Segals, B Perly, and H. Wroblewski, *Biochimica et Biophysica Acta*, **1235**, (2), 169-177 (1995).

53. S.H. White and W.C. Wimley, *Annual Review of Biophysics And Biomolecular Structure*, **28**, 319-365 (1999).

54. C. M. Deber and S. C. Li, *Biopolymers*, **37**, (5), 295-318 (1995).

55. C.J. Russell, T.E. Thorgeirsson, and Y.K. Shin, *Biochemistry*, **38**, 337-346 (1999).

56. L.Y. Shen, D. Bassolino, and T. Stouch, *Biophysical Journal*, **73**, 3-20 (1997).

57. C.M. Deber and N.K. Goto, *Nature Structural Biology*, **3**, 815-818 (1996).

58. L. A. Chung and T. E. Thompson, *Biochemistry*, **35**, (35), 11343-11354 (1996).

59. M. VanKann, J. Moller, H. Ringsdorf, and H. Hocker, *Journal of Colloid and Interface Science*, **178**, 241-250 (1996).

60. M. Gustafsson, *FEBS Letters*, **384**, 185-88 (1996).

61. S. Lee, H. Yoshitomi, M. Morikawa, S. Ando, and G. Sugihara, *Biopolymers*, **36**, 391-398 (1995).

62. T.B. Woolf, *Biophysical Journal*, **74**, 115-131 (1998).

63. M. A. Roseman, *Biophysical Journal*, **53**, 852-860 (1988).

64. S.H. White and W.C. Wimley, *Biochimica et Biophysica Acta-Reviews on Biomembranes*, **1376**, 339-352 (1998).

65. R.J. Webb, J.M. East, R.P. Sharma, and A.G. Lee, *Biochemistry*, **37**, 673-679 (1998).

66. J.A. Killian, *Biochimica et Biophysica Acta-Reviews on Biomembranes*, **1376**, 401-416 (1998).

67. L.R. Forrest, D.P. Tieleman, and M.S.P. Sansom, *Biophysical Journal*, **76**, 1886-1896 (1999).

68. L.A. Meijer, F.A.M. Leermakers, and J. Lyklema, *Journal of Chemical Physics*, **110**, 6560-6579 (1999).

69. E. Lopez-Hernandez and L. Serrano, *Proteins*, **22**, 340-349 (1995).

70. P. Ducarme, M. Rahman, and R. Brasseur, *Proteins-structure Function And Genetics*, **30**, 357-371 (1998).

71. J. A. Marquesee and K. A. Dill, *Journal Chemical Physics*, **85**, 434-44 (1986).

72. S. Marcelja, *Biochimica et Biophysica Acta*, **367**, 165-176 (1974).

73. M. Rooman and D. Gilis, *European Journal of Biochemistry*, **254**, 135-143 (1998).

74. B. Low, H. Reston, A. Sato, L. Searl, A. Rudco, and J. Richardson, *Proceedings National Academy Science*, **73**, 2991-2994 (1976).

75. W. R. Fiori and G. L. Millhauser, *Biopolymers*, **37**, (6), 421 (1995).

76. O. Ptitsyn, *J. Mol. Biol.*, **42**, 501-510 (1969).

77. J. Deisenhofer, O. Epp, K. Miki, R. Huber, and H. Michel, *Nature*, **318**, 618-624 (1985).

78. W. R. Fiori, S. M. Miick, and G. L. Millhauser, *Biochemistry*, **32**, 11957-11962 (1993).

79. W. S. Young and C. L. Brooks, *Journal Molecular Biology*, **259**, 560-572 (1996).

80. C. Kovacs, *Proceedings National Society Science*, **96**, 7901-7915 (1999).

81. B. H. Zimm and J. Bragg, *Journal Chemical Physics*, **31**, 526-535 (1959).

82. A. Yang and B. Honig, *Journal Molecular Biology*, **252**, 351-365 (1995).

83. M. L. Smythe, S. E. Huston, and G. R. Marshall, *Journal of the American Chemical Society*, **117**, 5445-5452 (1995).

84. S. S. Sung, *Journal Theoretical Biology*, **173**, 389-400 (1995).

85. J.A.W. Heymann and S. Subramaniam, *Proceedings of The National Academy of Sciences of The United States of America*, **94**, 4966-4971 (1997).

86. C. Tanford, *The Hydorophobic Effect: Formation of Micelles and Biological Memranes,* Wiley-Interscience, New York (1980).

87. A. G. Palmer, *Current Opinion in Biotechnology*, **4**, (4), 385-391 (1993).

88. A. Seelig and J. Seelig, *Biochemistry*, **13**, 4839-4845 (1974).

89. X. Peng and J. Jonas, *Biochemistry*, **31**, 6383-6390 (1992).

90. D. J. Siminovitch, P. T. T. Wong, and H. H. Mantsch, *Biochimica et Biophysica Acta*, **900**, 163-167 (1987).

91. H. L. Casal, *Journal of Physical Chemistry*, **93**, 4328-4330 (1989).

92. R. J. Cherry, *Biophysical Journal*, **64**, 1651-1652 (1993).

93. M. Milik, J. Skolnick, and A. Kolinski, *Journal of Physical Chemistry*, **96**, 4015 (1992).

94. S. J. Marrink, M. Berkowitz, and H. J. C. Berendsen, *Langmuir*, **9**, 3122-3131 (1993).

95. S. J. Marrink and H. J. C. Berendsen, *Journal of Physical Chemistry*, **98**, 4155-4168 (1994).

96. K. V. Damodaran and K. M. Merz, *Langmuir*, **9**, 1179-1183 (1993).

97. R. E. Jacobs and S. H. White, *Biochemistry*, **25**, 2605-2612 (1986).

98. J. Seelig, *Quarterly Review of Biophysics*, **10**, 2605-2612 (1977).

99. G. Buldt, H. U. Gally, J. Seelig, and G. Zaccai, *Journal of Molecular Biology*, **134**, 673-691 (1979).

100. M. J. Janiak, D. M. Small, and G. G. Shipley, *Journal Biological Chemistry*, **254**, 6068-6078 (1979).

101. K. V. Damodaran, K. M. Merz, and B. P. Gaber, *Biochemistry*, **31**, 7656-7664 (1992).

102. G. Cevc, *Chemistry and Physics of Lipids*, **64**, 163-186 (1993).

103. J-P. Douliez, A. Leonard, and E. J. Dufourc, *Biophysical Journal*, **68**, (5), 1727-1739 (1995).

104. A. J. Robinson, W. G. Richards, P. J. Thomas, and M. M. Hann, *Biophysical Journal*, **67**, 2345-2354 (1994).

105. H. Heller, M. Schaefer, and K. Schulten, *Journal of Physical Chemistry*, **97**, 8343-8360 (1993).

106. F. Zhou and K. Schulten, *Journal of Physical Chemistry*, **99**, 2194-2207 (1995).

107. J. N. Isrealuchvili, S. Marcelja, and S. Horn, *Quarterly Review of Biophysics*, **13**, 121-200 (1980).

108. N. Albon, *Journal Physical Chemistry*, **89**, 3143-5 (1985).

109. R. Mendelsohn, M. A. Davies, J. W. Brauner, H. F. Schuster, and R. A. Dluhy, *Biochemistry*, **28**, (22), 8934-8939 (1989).

110. K. Jorgensen, A. K. Linger, M. Braimen, and R. L. Biltonen, *Journal of Physical Chemistry*, **100**, 2766-69 (1996).

111. K. Tu, D. J. Tobias, J. K. Blasie, and M. L. Klein, *Biophysical Journal*, **70**,

595-608 (1996).

112. S. E. Feller, Y. Zhang, and R. W. Pastor, *Journal of Chemical Physics*, **103**, (23), 10267-76 (1995).

113. S. W. Chiu, M. Clarke, V. Balaija, S. Subramaniam, H. L. Scott, and E. Jakobsson, *Biophysical Journal*, **69**, 1230-45 (1995).

114. J. F. Nagle, *Biophysical Journal*, **64**, 1476-1481 (1993).

115. D.P. Tieleman and H.J.C. Berendsen, *Journal of Chemical Physics*, **105**, (11), 4871-80 (1996).

116. R. W. Pastor, R. M. Venable, and M. Karplus, *Proceedings of the National Academy of Sciences of the United States Of America*, **88**, 892-896 (1991).

117. M.R.R. de Planque, D.V. Greathouse, R.E. Koeppe, H. Schafer, D. Marsh, and J.A. Killian, *Biochemistry*, **37**, 9333-9345 (1998).

118. A. Johannson, E. A. Smith, and J. A. Metcalfe, *Biochimica et Biophysica Acta*, **641**, 416-421 (1981).

119. M. D. Reboiras and D. Marsh, *Biochimica et Biophysica Acta*, **1063**, 259-264 (1991).

3. Methods

3.1. Potential Energy Force Field

In empirical calculations a molecule is represented in terms of its constituent atom types e.g. Csp^2, Csp^3 rather than explicitly calculating the electronic configurations as in *Ab Initio* Quantum Mechanics calculations. These atomic representations are described by a potential energy function and in this study the **Valence Force Field, V.F.F.,**[1] potential function was used.

In the *V.F.F.* the potential energy,*V* of the system is represented by an analytical function of the internal coordinates and non-bonded interactions. The Valence Force Field includes the following terms:

$$V = E_{bond} + E_{angle} + E_{torsion} + E_{oop} + E_{cross} + E_{vdW} + E_{electrostatic} \qquad (2.1)$$

The bond strain energy is represented by an exponential 'Morse' function.

$$E_{bond} = \sum_b \left(D_b[1 - \exp(-\alpha(b - b_0))]^2 - D_b \right)$$

b and b_0 are the observed and equilibrium bond lengths. The parameter D_b is the depth of the potential well and parameter α is the well shape.

$$E_{bond} = \tfrac{1}{2} \sum_b K_b'(b - b_0)^2$$

A harmonic representation of the bond energy, used with highly strained systems, where k'_b is the force constant for the bond.

The angle strain is represented by a harmonic function.

$$E_{angle} = \tfrac{1}{2} \sum_\theta K_\theta(\theta - \theta_0)^2$$

K_θ is the force constant for angle bending. θ and θ_0 are the observed and equilibrium angles.

The torsional energy is represented by parameters producing the required number of minima for the three states of hybridisation.

$$E_{torsion} = \tfrac{1}{2}\sum_{\phi} K_{\phi}[1 + T\cos(n\phi)]$$

K_{ϕ} is the force constant for torsional deformations and T takes values 1 or -1. n is 1, 2, 3 depending on the hybridisation of the atoms. ϕ is the observed torsion angle.

Sp^2 hybridised atoms require an out-of-plane term.

$$E_{oop} = \tfrac{1}{2}\sum_{\chi} K_{\chi}\chi^2$$

χ is the out of plane angle, K_{χ} is the force constant for out of plane deformations.

Other additional terms describe the coupling between these terms, i.e. the energy required to deform one internal coordinate depends on the current value of another internal coordinate.

$$E_{cross} = E_{bb'} + E_{\theta\theta'} + E_{b\theta} + E_{\phi\theta\theta'} + E_{\chi\chi'} \quad \cdots \quad \text{Cross Terms} \tag{2.2}$$

$$E_{bb'} = \sum_{b}\sum_{b'} F_{bb'}(b - b_0)(b' - b_0')$$

$$E_{\theta\theta'} = \sum_{\theta}\sum_{\theta'} F_{\theta\theta'}(\theta - \theta_0)(\theta' - \theta_0')$$

$$E_{b\theta} = \sum_{b}\sum_{\theta} F_{b\theta}(b - b_0)(\theta - \theta_0)$$

$$E_{\phi\theta\theta'} = \sum_{\phi} F_{\phi\theta\theta'}\cos\phi(\theta - \theta_0)(\theta' - \theta_0')$$

$$E_{\chi\chi'} = \sum_{\chi}\sum_{\chi'} F_{\chi\chi'}\chi\chi'$$

The non-bonded interactions are described by a Lennard-Jones 6-12 term.

$$E_{vdW} = \sum_{nb} \varepsilon[(r^*/r_{ij})^{12} - 2(r^*/r_{ij})^6]$$

r^* is the equilibrium bond distance for atoms i and j, ε the minimum energy value at r^*.

The electrostatic interactions are represented using a Coulombic term based on point charges.

$$E_{electrostatic} = \sum_{nb}(q_i q_j)/\varepsilon r_{ij}$$

q_i and q_j are the charges of the two non-bonded atoms in the non-bonded pair. ε the dielectric takes the value 1.0, r_{ij} is the distance between the two non-bonded atoms. Values for q_n the atomic charges have been derived from *ab initio* quantum mechanics calculations and charge fitting calculations.[2-5]

The importance of the quality of the form and parameterisation of the force field has been shown in several studies.[6-10] The potential energy function and parameters for lecithin molecules have been derived from MD and M.M. simulations of dilaurylglycerol and glycerol phosphatidylcholine.[4] These have being developed using the same implementation of the force field[11] and hence are compatible with parameters from the standard V.F.F. potential library. Atomic charges were also taken from this study, derived from a linear least squares fit to the value of the electrostatic potentials of the molecules in the unit cell and those molecules surrounding it.

The first and second derivatives of the potential energy function with respect to the Cartesian coordinates can be obtained analytically.[12] The derivatives of the energy with respect to the internals can be derived directly from (Eq. 2.1). The derivatives of the internals with respect to the Cartesian can be derived from the equations defining the internals in terms of the Cartesian coordinates, obtained from geometry considerations.

The first derivatives are given by:

$$\frac{\partial V}{\partial x_i} = \sum_a \frac{\partial V}{\partial l_a} \frac{\partial l_a}{\partial x_i} \tag{2.3}$$

The second derivatives are given by:

$$\frac{\partial^2 V}{\partial x_i x_j} = \sum_a \sum_b \frac{\partial^2 V}{\partial l_a \partial l_b} \frac{\partial l_b}{\partial x_j} \frac{\partial l_a}{\partial x_i} + \sum_a \frac{\partial V}{\partial l_a} \frac{\partial^2 l_{\mathring{A}}}{\partial x_i \partial x_j} \tag{2.4}$$

l_a and l_b are the internal coordinates (bonds, angles etc.), the functions are summed over all the internals a and b.

3.2. Energy Minimisation

The energy of the particular system is minimised by solving the equation:

$$\partial V/\partial x_i = 0 \qquad i = 1, 2, 3, \cdots, 3n \tag{2.5}$$

x_i are the Cartesian coordinates of the system and n is the number of atoms.

The *steepest descents* algorithm was used in this study to solve the equation 2.5. This form of minimiser uses the the gradient of the energy surface, i.e. the first derivative, to determine the search direction.

$$X_n = X_{n-1} + \lambda_n S_n \qquad s_n = -g_n/|g_n| \tag{2.6}$$

x_n is the coordinate vector at iteration n, s_n is the direction of the step as determined by the energy gradient vector, g_n and the step size, λ_n.

If the energy is increased for a step, i.e. the gradient is positive, the step size is reduced and the previous step is repeated. If the energy decreases then the step direction is updated and the step size is increased. This process is repeated until the required minimum energy is achieved, i.e. the relaxed structure, (Eq. 2.5).

The steepest descent algorithm is a powerful tool to relax a strained initial structure or a system with clashes, as it requires less evaluation of the derivatives to achieve the required low energy. But as the minimum is approached it becomes

inefficient and has a slow speed of convergence.

3.3. Molecular Dynamics

Classical mechanics Lagrange equation of motion describes the motion of a particle system;[12]

$$\frac{d}{dt}\frac{\partial K}{\partial \dot{q}_i} + \frac{\partial V}{\partial q_i} = 0 \qquad i = 1, 2, 3, \cdots, 3n \tag{2.7}$$

Where K and V are the kinetic and potential energy of the system, and $\dot{q}_i$ and q_i are the velocities and coordinates.

In a system of Cartesian coordinates, x_i, (and corresponding velocities and accelerations, $\dot{x}_i$ and $\ddot{x}_i$, respectively) the kinetic energy is given by :

$$K = \tfrac{1}{2}\sum_{i=1}^{3n} m_i \dot{x}_i^2 \tag{2.8}$$

and the forces are defined by:

$$F_i = -\frac{\partial V}{\partial x_i} \tag{2.9}$$

Using (Eq. 2.9) and (Eq. 2.8) in the Lagrange equation, (Eq. 2.7), we obtain the familiar form of Newton's equations of motion:

$$F_i = m_i \ddot{x}_i \qquad i = 1, 2, 3, \ldots, 3n \tag{2.10}$$

In molecular dynamics, *MD,* these classical equations are solved iteratively. The force on an atom *i* is calculated directly from the potential energy derivative ∂V, with respect to the coordinate x_i (Eq. 2.9).

$$-\frac{\partial V}{\partial x} = m\ddot{x} \tag{2.11}$$

Hence the accelerations can be obtained from these forces (Eq. 2.10).

The motion of an atom can be expressed in terms of a standard Taylor series, where a

small time step (Δt) is applied to generate the next coordinate.

$$x(t+\Delta t) = x(t) + \dot{x}\Delta t + \ddot{x}\frac{\Delta t^2}{2!} + \cdots \tag{2.12}$$

In the application of the Taylor Expansion to the present coordinates, in order to produce a new set of coordinates, the higher order terms are ignored which results in a small error. The numerical solutions rely on knowing the position x(t), velocity $\dot{x}$, acceleration $\ddot{x}$. This is the underlying numerical approximation of all the dynamics algorithms.

In this study the specific algorithm used is the *Velocity* Verlet algorithm as it calculates the velocities and coordinates at each time step.

The *Velocity Verlet* algorithm.[13]

$$x(t+\Delta t) = x(t) + v(t).\Delta t + \tfrac{1}{2}a(t).\Delta t^2 \tag{2.13}$$

$$v(t+\tfrac{1}{2}\Delta t) = v(t) + \tfrac{1}{2}a(t).\Delta t \tag{2.14}$$

$$a(t+\Delta t) = -\partial V/\partial x.\ 1/m \tag{2.15}$$

$$v(t+\Delta t) = v(t+\tfrac{1}{2}\Delta t) + \tfrac{1}{2}a(t+\Delta t).\Delta t \tag{2.16}$$

Hence the velocities, accelerations and coordinates at each time step are known.

The initial velocities are assigned to a Maxwell-Boltzmann (Eqn.2.27) distribution at the desired temperature. A time step of 1 fs is used because the highest frequency motion is the O-H stretch of water (3700 cm^{-1} or 11.09×10^{15} s^{-1}).

3.4. Statistical Ensembles

The above described MD algorithms sample in the microcanonical (NVE) ensemble with a constant number of particles, volume and the total energy. The physiological membrane is measured at NPT. These conditions can be approached with modifications to the dynamics algorithm described.

3.4.1. NVT Conditions

The temperature of a system is related to the kinetic energy and the atomic velocities via the kinetic theory of gases and is the average kinetic energy per degree of freedom:

$$U = 3/2.\, k_b T \tag{2.17}$$

U is the kinetic energy, k_b is the Boltzmann constant

The NVT ensemble is simulated by coupling the NVE system to a temperature bath.[14, 15] The Berendsen scaling factor,[14] λ, is applied to the velocities,

$$\lambda^2 = 1 + \frac{\Delta t}{\tau}\left(\frac{T_0}{T} - 1\right) \tag{2.18}$$

Δt is the time step size, τ is the relaxation time, T_0 is the target temperature, and T is the instantaneous temperature.

For the Velocity Verlet integration algorithm the "ad hoc" rescaling algorithm[15] is used.

$$r(t + \Delta t) = r(t) + \lambda \Delta t v(t) + (2 - \lambda)\frac{\Delta t^2}{2} a(t) \tag{2.19}$$

where λ is Berendsen's scaling factor (eqn. 2.18), τ_t is the coupling time constant, T_0 is the target temperature and T is the actual temperature.

All the simulations in this study were performed using the temperature bath.

3.4.2. NPE Conditions

The pressure, essentially, can be thought of as the energy per volume and can be defined by two contributions, one from the ideal and the other from the non-ideal behaviour of the system. The ideal contribution is so called as it is the only contribution to the pressure in an ideal gas (in which the potential energy is zero, i.e. from kinetic motion of particles).

$$(PV)_{ideal} = Nk_bT = 1/3 < \sum_{i=1}^{N} |p_i|^2/m_i > \quad (2.20)$$

N is the number of atoms, k_b is the Boltzmann constant, p_i is the momentum and m_i is the mass of atom i.

The non-ideal component, the virial ω, comes from the potential energy of the system.

$$\omega = -1/3 \sum_{i=1}^{N} r_i \nabla_{r_i} V = 1/3 \sum_{i=1}^{N} r_i f_i \quad (2.21)$$

r_i is the coordinate of atom i, $\nabla_{r_i} V$ is the derivative of the potential energy V.

Hence for intermolecular interactions:

$$PV = Nk_bT + < \omega > \quad (2.22)$$

The instantaneous pressure P_i is

$$P_i = Nk_bT_i + \omega/V \quad (2.23)$$

T_i is the instantaneous temperature as calculated from the velocities. *V* is the volume of the unit cell.

The NPE conditions are simulated by coupling the NVE system to an external pressure using the following expression.[14]

$$\delta P/\delta t = (P - P_i)/t_p \quad (2.24)$$

P is the target pressure and t_p is the pressure scaling time constant. At each step the volume of the box is scaled by χ, and the coordinates of the molecular centres of mass by $\chi^{1/3}$.

$$x_i' = \chi^{1/3} x_i \quad (2.25)$$

where

$$\chi = 1 - \beta_T \frac{\delta t}{t_p} (P - P_i) \quad (2.26)$$

β_T is the isothermal compressibility of the system, δt the time step of the simulation, t_p takes the value 0.5ps.

The volume of the box is changed by scaling the unit cell lengths by the pressure tensor matrix diagonals.[16]

In the simulations included in this thesis only isotropic scaling of the Unit Cell Vectors were used. In addition, in the VFF the scaling can be limited to an individual cell dimension or to a combination of two or all of the three (a,b,c). Unit Cell scaling was used during the equilibration phase of the simulations to better control the box size changes.

3.4.3. Isobaric-Isothermal (NPT) Conditions

One may also sample the molecular dynamics of the system using both of the above baths. Hence the system is both isothermal and isobaric. This was used during the equilibration phase of the simulation.

3.5. Periodic Boundary Conditions

We use Periodic boundary conditions, *P.B.C.*, to reduce surface or edge effects.[16] The system coordinates are defined within a cubic box known as "the unit cell", this box is then replicated in 3-dimensions throughout space in an infinite crystal lattice.

The system works to keep the density of the box constant, so, if in a simulation a molecule leaves the unit cell, a *"basic"* molecule, via its centre of mass, then a replicate molecule, a *"ghost"* (an identical molecule from an adjacent box), enters the unit cell from the opposite face. During a simulation under *periodic boundary conditions* (P.B.C.) molecules move in and out of the unit cell and these density maintenance translations are known as "swapping". In P.B.C. the basic molecules interact with both the other basic molecules and the ghost molecules, whilst the ghost molecules only interact with the basic molecules. A cut-off is used to limit the non-bonded calculations to a finite system.

References

1. P. Dauber, M. Goodman, A. T. Hagler, D. J. Osguthorpe, R. Sharon, and P. S. Stern, *Proceedins of the ACS Symposium on Supercomputers in Chemistry*, **173**, 161 (1981).
2. B. H. Besler, K. M. Merz, and P. A. Kollmann, *Journal Computational Chemistry*, **11**, (4), 431-439 (1990).
3. D. E. Williams and T. R. Stouch, *Journal of Computational Chemistry*, **14**, (9), 1066-1076 (1993).
4. T. R. Stouch, K. B. Ward, A. Altieri, and A. T. Hagler, *Journal of Computational Chemistry*, **12**, 1033-1046 (1991).
5. P. S. Charifson, R. G. Hiskey, and L. G. Pedersen, *Journal Computational Chemistry*, **11**, (10), 1181-1186 (1990).
6. M. Billeter, A. E. Howard, I. E. Kultz, and P. A. Kollmann, *Journal American Chemical Society*, **110**, 8385 (1988).
7. A. T. Hagler, S. Lifson, and P. Dauber, *Journal American Chemical Society*, **101**, 5122 (1979).
8. J. R. Mapel, U. Dinur, and A. T. Hagler, *Proceedings of the National Academy of Sciences*, **85**, 5350 (1988).
9. A. T. Hagler, P. S. Stern, S. Lifson, and S. Ariel, *Journal American Chemical Society*, **101**, 813 (1979).
10. P. Dauber-Osguthorpe, V. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, *Protein : Structure Function and Genetics*, **4**, 31-47 (1988).
11. R. G. Snyder and J. H. Schachtschneider, *Spectrochim. Acta*, **19**, 116 (1963).
12. P. Dauber-Osguthorpe. Ph.D. Thesis Bath, 1990
13. L. Verlet, *Phy. Rev.*, **159**, 98 (1967).

14. H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola, and J. R. Haak, *Journal of Chemical Physics*, **81**, 3684-3690 (1984).

15. D. Brown and J. H. R. Clarke, *Molecular Physics*, **51**, (5) (1984).

16. M. P. Allen and D. J. Tildesley, *Computational Simulations of Liquids,* Clarenden Press, Oxford (1987).

4. Software and Analysis Tools

4.1. The V.F.F.

In this study the program *V.F.F.*[1] was used to perform both the minimisation and molecular dynamics. In order to perform the calculations involved, the program requires the input of certain data:

(1) Simulation File - a file that the program uses to govern the simulation details, containing information about the required physical conditions, geometric constraints and restraints, plus the simulation timing details.

(3) Parameter File - the contents include all the V.F.F. potential parameters,

(2) Residue Library - contains the atomic connectivity details and the bond hybridisation details,

(4) Cartesian Coordinate file - containing the 3-dimensional coordinates of all atoms the system,

This information is used by the program to attain both the required local minima and to generate molecular dynamics trajectories of properties of the system, such as coordinates, velocities and energy components.

4.2. Insight

Insight [2] is a molecular graphics package that allows the building of a molecule on an atomistic basis or by the joining of fragments that the Insight library contains, in order to create a cartesian coordinate file and a molecular data file containing connectivity information. Insight will display the molecule in both 2D and 3D, whilst allowing the user to manipulate the molecule, not just in the translation and rotational motion of the molecule, residue or fragment, but also via alterations in the internal geometry, e.g. the valence and dihedral angles, the bond distances and hybridisation. Insight also allows the mutation of atoms, groups, amino acid residues and molecules.

It allows the viewing of molecular dynamics trajectories and the calculation of a small number of system properties. It can also solvate a molecule using the internally stored Monte Carlo generated box of water molecules at $\rho = 1$ g cm^{-3}, whilst avoiding Van der Waal clashes within the solvated system. The *Cell* command can be used to solvate the system using *Periodic Boundary Condition* parameters in order to produce a solvated periodic unit cell as opposed to a "Droplet" solvation.

4.3. Elim_H2O

This is a program[3] that allows for the elimination of water molecules from the hydrophobic region of a phospholipid bilayers. When a molecule is solvated by Insight, for example, water molecules are placed in all available space of a large enough volume, so the removal of water molecules from the hydrophobic regions is required in the modelling of phospholipid bilayers.

4.4. SOAK

The program *SOAK*[4] was used to solvate the poly-alanine helix by replicating a prepared box of liquid water. The program takes a molecular system from an input coordinate file, reorientates the molecular system onto the required coordinate axis and solvates the system to predetermined dimension using an input solvent box. A Van der Waal clash check is applied to allow for the removal of both solvent-solute clashes and boundary solvent- solvent clashes up to a predetermined threshold. This is done so as to retain the bulk density near to that of the initial value. [The remaining clashes being reduced when a minimiser is applied to the system].

4.5. MOLEDT

This is a molecule building program[2] for peptides and proteins, that takes user input parameters and data on the required primary, secondary and tertiary structure of a peptide or protein molecule and outputs a cartesian coordinate file and molecular

data file containing connectivity and bond order information.

4.6. FOCUS

FOCUS, Finally One Can Understand Simulations[5] , is a program that analyses the output trajectory from a molecular dynamics simulation. *FOCUS* can calculate properties at each time-step and time averaged properties directly from the molecular dynamics trajectory or it can reformat the coordinates, atomic velocities and energies from a format of all properties at a given time-step to a format of a given property over all time-steps. The program requires a molecular dynamics trajectory "history" file and a control file.

The focus analysis tools used in this study are outlined below:

4.6.1. Time Average and Standard Deviation

The statistical average of property, p ;

$$< p > = \frac{1}{N} \sum p(t_k)$$

where, <p> is the time average of p and $p(t_k)$ is the property value at time k.

The Standard Deviation, σ, of property, p ;

$$\sigma = \left(< p^2 > - < p >^2 \right)^2$$

where, $< p^2 >$ time average of the property squared.

4.6.2. Distance

The distance between any 2 atomic coordinates can be calculated, e.g. a bond length or a non-bond distance. A list of bonded and non-bonded atom pairs is input and the distance between the bonded or non-bonded atom pairs is calculated across the trajectory or a required section of the trajectory.

4.6.3. Angles

Bond angles can be calculated between three input bonded atoms, a pseudo-valence angle can also be calculated between three non-bonded atoms and a torsion or dihedral angle between four bonded atoms can also be calculated, as well as a pseudo-torsion angle. A list of bonded and non-bonded angles and dihedral angles are input into FOCUS and the angles are calculated across the trajectory. The torsion angle data can be used to determine the distribution of trans and gauche (t/g) rotational isomers across the trajectory for a particular angle or the average value for a particular molecule.

4.6.4. Radial Distribution Function

The pair radial distribution function,[6] g(r), of a particular atom or a sub-section of atoms can be calculated in a 2 dimensional plane with respect to a specified atom or atoms (e.g. the oxygen atom of a solvating water). For all pair separations, the distance (r) from the reference atom or atoms are calculated and placed into bins. For P.B.C. systems distances between basic-ghost atoms are also included. The bin dimensions are determined by the resolution required, δr and each bin is normalised.

$$n(i) = \sum_{1} n_r + \delta r$$

where, n(i) is the average number of distances in the range r,r+δ_r.

The ensemble average is calculated and then the expected density is used to normalise the radial distribution.

$$n^{nrm} = \frac{4\pi\rho}{3}\left[(r+\delta r)^3 - r^3\right]$$

where, n^{nrm} = number of entries found in radial segment (r+δ r) - r, and

ρ = bulk density of particle.

$$g(r+\delta r) = \frac{n(i)}{n^{nrm}(i)}$$

4.6.5. Density Profile

This feature allows the calculation of the density of a set of atoms and/or molecules with respect to a specific coordinate axis. The masses of the atom set are summed across a series of bins resulting in a density profile for the unit cell. In the pressure bath simulations the unit cell will vary in volume, so a *Unit Cell Expansion Factor* is applied in order to scale the selected axis to maintain the values within the range for the whole simulation time period.

4.6.6. Segmental Order Parameter, SOP

The segmental order parameters for methylene groups in the acyl chains of the lipid molecules is calculated as follows.

A tensor, S_{ij} is calculated: $$S_{ij} = \frac{1}{2} < 3\cos\theta_i \cos\theta_j - \delta_{ij} >$$

where i,j = [x,y,z], δ_{ij} is 0 if i=j or 1 if i≠j and θ_i is the angle subtended between that axis and the fixed system coordinate axis defined as the bilayer normal (the y-axis in our bilayer systems).

The three molecular axes of the methylene groups are defined thus:

x-axis = H-H vector

y-axis = bisectrix of the H-C-H angle

z-axis = the vector perpendicular to the H-C-H plane

This defines the methylene groups average orientation relative to a long molecular axis (the bilayer normal) to experimental results from ^{2}H NMR spectroscopy, (C-D bonds). The order parameter (S_{cd}) calculated below can be compared to experimental ^{2}H NMR spectroscopy results (C-D bonds).

$$S_{cd} = 2/3\ S_{xx} + 1/3\ S_{yy}$$

In an all trans chain with the chain aligned along the bilayer normal then the segmental order parameter for each methylene carbon is 1, whilst in a completely random

chain, i.e. totally disordered, it is 0.

4.6.7. Root Mean Square Deviation, RMSD

In this feature the coordinate structure from a molecular dynamics trajectory is superimposed onto an input template coordinate file containing either the initial structure or a crystal structure, and via rigid body translations and rotations the best fit is produced. The Root Mean Square Deviation is calculated between the whole structure or sub-sections of the structure, thus, giving a quantitative measure of the structural similarities.

This procedure can be used to compare not only an external reference file, e.g. the initial structure, but also a previous designated structure from the molecular dynamics trajectory.

4.6.8. Thermodynamic Data

A molecular dynamics trajectory contains information about the simulation condition at each time step which can be output as either time dependent properties or time averaged properties. The physical conditions include - unit cell volume, angles and lengths, temperature and pressure and the energies (kinetic, potential, total) and the energy components e.g. E_b, E_{theta}, E_ϕ, e.t.c..

4.6.8.1. Atomic and Molecular Temperature

The kinetic theory of gases can be used to calculate the temperature of an atom, a group of atoms or a molecule from their velocities :

$$K.E. = \frac{3}{2} k_b T$$

$$K.E. = \frac{1}{2} mv^2$$

$$T = \sum_{i=1}^{N} \frac{m_i v_i^2}{3k_b}$$

where, m_i = mass of atom i, v_i = velocity of atom i.

This process allows the calculation of the temperature of a set of atoms. Therefore, it is possible to investigate the partitioning of the temperature among the atoms and its resultant effects and contribution to the dynamic properties of the system.

4.6.8.2. Partial Molecular Energies

Coordinate files were dumped out from the MD trajectory using FOCUS at regular periods during the simulation. The energy of each individual coordinate file was calculated (using the V.F.F.) and partitioned into the inter-molecular and intra-molecular partial atomic non-bond energies of the helix atoms. The partial non-bond energy for atom *i* is the non-bond interaction energy between atoms *i* and *j*, divided equally between the two participating atoms, and summed over atoms *j* in the system,

$$E_i = \tfrac{1}{2} \sum_j V_{r_{ij}}$$

These partial atomic non-bond energies were partitioned into the three component parts of the non-bond interactions, the repulsive and dispersive Lennard-Jones terms and the coulombic terms. The manipulation of these energies into residue based energies allows the characterisation of the helix structure and the solvent environment induced effects.

4.7. Definitions

The tables below contain the atom number definition of the structural internal angles and the torsion angles, as per the DMPC molecule on Page 75, and the residue number of the atom in the helix. These definitions will be used in the results chapters to come.

Table 4.1

Helix Structural Internals			
Lengths		Angles	
d_1	CA1-CA32	θ_1	CA1-N17-CA32
d_2	CA1-N17	θ_2	CA1-N11-CA32
d_3	N17-CA32	θ_3	CA1-N23-CA32
d_4	CA1-N11		
d_5	N11-CA32		
d_6	CA1-N23		
d_7	N23-CA32		

Table 4.2

Lipid Structural Internals			
Lengths		Angles	
d_1	N-P	θ_1	N-P-C6
d_2	N-C6	θ_2	C21-C6-C35
d_3	C6-C21	θ_3	N-C6-C35
d_4	C6-C35	θ_4	N-C6-C21
d_5	C9-C21	θ_5	C6-C23-C35
d_6	C23-C35	θ_6	C6-C9-C21
d_7	C21-C35		
d_8	P-P		

Table 4.3

Lipid Torsion Angles			
No.	Torsion	No.	Torsion
1	C - N - C3 - C4	22	C16 -C17 -C18 -C19
2	C1 - N - C3 - C4	23	C17 -C18 -C19 -C20
3	C2 - N - C3 - C4	24	C18 -C19 -C20 -C21
4	N - C3 - C4 - O	25	O3 - C5 - C6 - O6
5	C3 - C4 - O - P	26	C5 - C6 - O6 -C22
6	C4 - O - P - O3	27	C6 - O6 -C22 -C23
7	O - P - O3 - C5	28	O6 -C22 -C23 -C24
8	P - O3 - C5 - C6	29	C22 -C23 -C24 -C25
9	O3 - C5 - C6 - C7	30	C23 -C24 -C25 -C26
10	C5 - C6 - C7 - O4	31	C24 -C25 -C26 -C27
11	C6 - C7 - O4 - C8	32	C25 -C26 -C27 -C28
12	C7 - O4 - C8 - C9	33	C26 -C27 -C28 -C29
13	O4 - C8 - C9 -C10	34	C27 -C28 -C29 -C30
14	C8 - C9 -C10 -C11	35	C28 -C29 -C30 -C31
15	C9 -C10 -C11 -C12	36	C29 -C30 -C31 -C32
16	C10 -C11 -C12 -C13	37	C30 -C31 -C32 -C33
17	C11 -C12 -C13 -C14	38	C31 -C32 -C33 -C34
18	C12 -C13 -C14 -C15	39	C32 -C33 -C34 -C35
19	C13 -C14 -C15 -C16	40	C22 - O6 - C6 - C7
20	C14 -C15 -C16 -C17	41	O6 - C6 - C7 - O4
21	C15 -C16 -C17 -C18		

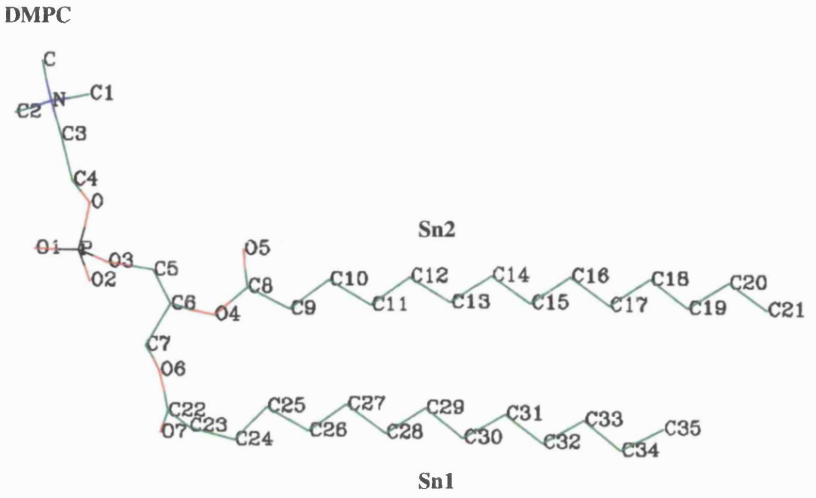


Fig. 4.1 - Schematic Structures of the DMPC Molecule

Common Abbreviations

The abbreviations listed below are used in the following results section's text and in the accompanying figures or tables.

System Sub Groups:

Sys=system, PC=phosphatidyl choline group, GL=glycerol moiety, Sn1/Sn2=acyl chains, Wtr=inter-bilayer water

Thermodynamic Data:

KE=Kinetic Energy, PE=Potential Energy, Tot=Total Energy, EB=Bond Energy, ET=Theta Energy, EP=Phi Energy, DSP=VdW Dispersive Energy, REP=VdW Repulsive Energy, EST=Electrostatic Energy, TEMP=Temperature, PRESSRM=Molecular Pressure

References

1. P. Dauber, M. Goodman, A. T. Hagler, D. J. Osguthorpe, R. Sharon, and P. S. Stern, *Proceedins of the ACS Symposium on Supercomputers in Chemistry*, **173**, 161 (1981).
2. Biosym. Technologies Inc. San Diego, CA, USA
3. A. P. Lemon M.G.U. Bath, U.K.
4. P. J. Tollinton M.G.U., Bath, U.K.
5. D. J. Osguthorpe and P. Dauber-Osguthorpe, *Journal of Molecular Graphics*, **10**, (Sept.) (1992).
6. M. P. Allen and D. J. Tildesley, *Computational Simulations of Liquids,* Clarenden Press, Oxford (1987).

5. Pure Bilayer Simulations

5.1. Building The Bilayer

In this thesis two types of lipid packing have been used in order to the build model membranes, the cubic and hexagonal geometry. The different geometries were investigated using Molecular Dynamics (MD) techniques and their ability to reproduce experimental results and physical parameters analysed.

Hex - Hexagonally Packed Fully Hydrated DMPC Bilayer System

Sq - Cubic Packed Fully Hydrated DMPC Bilayer System

The arrangement of the two lipids on their different geometric lattices (Fig. 5.1) can be seen to differ and these differences may impact on the dynamic behaviour of the two bilayer systems during a MD simulation.

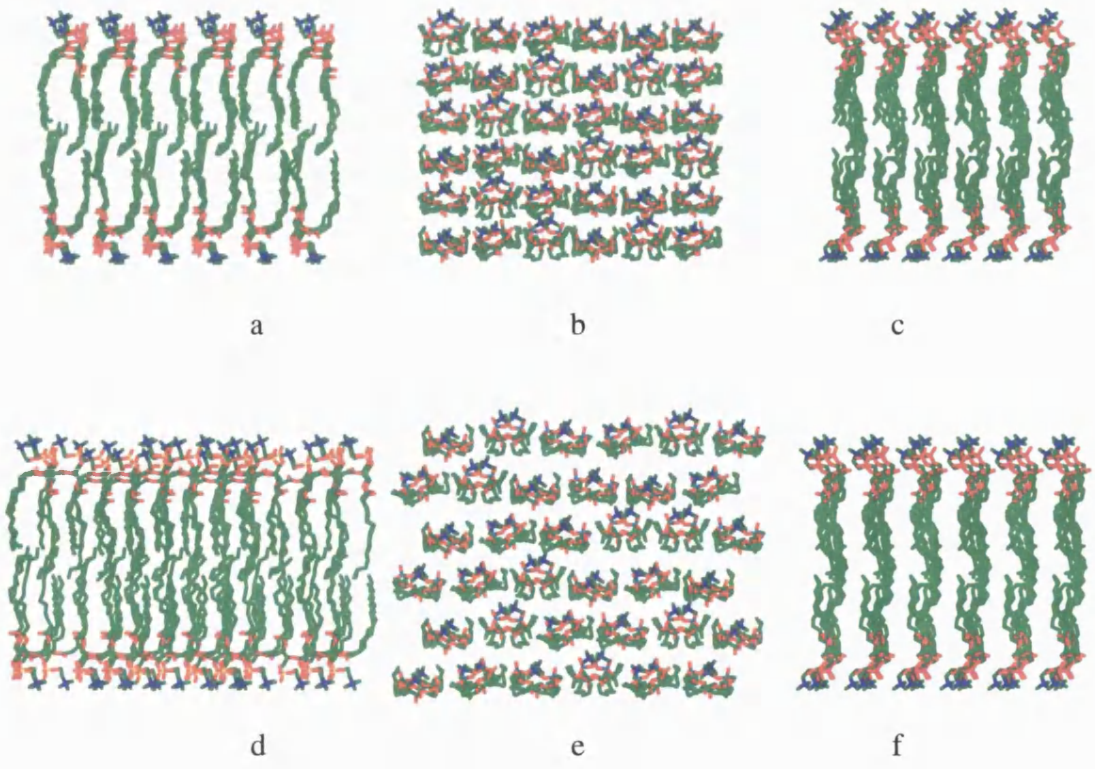


Fig. 5.1 Initial Coordinates Systems : Sq (a-c) Hex (d-f)

The cubic packed lipid system (Fig 5.1) shows the features highlighted below:

- The 'head on' view shows little inter-head group space (Fig. 5.1b)
- The 'side-on' view shows that the cubic geometry allows more lateral space between the lipids (Fig. 5.1a) than that seen for the head groups.
- The second 'side-on' view, 90° from the above view, shows that whilst there is little inter-lipid space, the "row & column" nature of a cubic system may direct the motion of the lipids towards cooperative modes (Fig 5.1c).

The Hexagonal structure has features dissimilar to the above which tend to suggest different motional tendencies and are highlighted below:

- The head group, 'head-on' view shows more interstitial space than the cubic system due to the 'diagonal' geometry (Fig. 5.1e).
- The side view clearly shows the head groups are wider spaced, to avoid clashes, thus the tail of the lipids possess more space than in the cubic system (Fig 5.1d).
- The second 'Side-on' view illustrates that the lipids are in an off-set lattice and gives the acyl chains a semi-random appearance (Fig. 5.1f)

5.1.1. The Building Protocol

The starting structure is a very important feature of any MD simulation because of the long equilibration time for lipid motion in a bilayer. Recently the process of building and equilibrating the bilayer structure has been refined, problems with cut-off induced temperature fluctuations, acyl chain rigidity and the use of a short cut-off, have been looked at in order to improve the efficiency and accuracy of simulations. To improve the rate of acyl chain equilibration a series of 4 lipids were built from acyl chains of lipid molecule molecules extracted from a simulation of a 5x5 DMPC hydrated bilayer. This simulation had been run for 200 ps and had been rejected for two reasons. Firstly, the final structure had a rigid acyl chain region as the application of a single temperature bath (Table 5.1) caused the acyl chains to be 'frozen' (explained later). The cooling of the acyl chains resulted in a reduced degree of chain

melting and a low gauche population for the acyl chain torsion angles of only ≈1 per chain, a value of 2 or 3 times more than this is observed experimentally. Secondly, the surface area per lipid head group (SA) and the cross-bilayer head group phosphorous-phosphorous distances (P-P) were not close to the experimentally observed values.

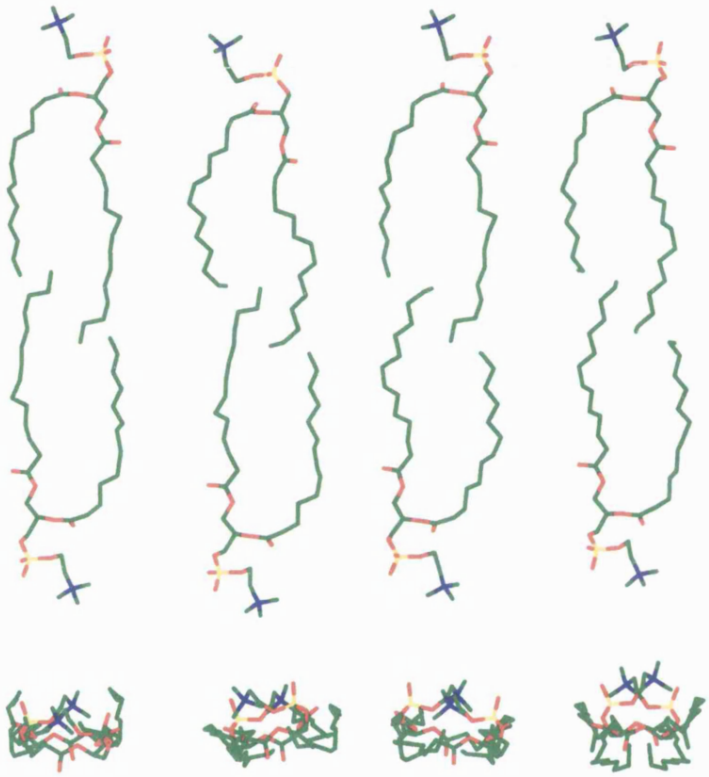


Fig. 5.2 - The Pre-equilibrated Lipid Pairs

The acyl chains were selected because they had a gauche torsion bond count per chain of 2 (close to the experimental value) and therefore would give a degree of pre-equilibration, reducing the equilibration time required.

The head groups were deleted from their acyl chains because they had developed a high temperature, 328° (Table 5.1), and, as the head groups are semi-crystalline, it was decided not to use these fluid structures in the present study. The original

crystalline lipid head group coordinates were used and the *Insight*[1] graphics suite used to align the selected pre-equilibrated acyl chains with the head group glycerol moiety. The lipids were also constructed to obtain a structure that facilitates packing onto the two geometric lattices. The new lipid structures were then used to construct pairs of lipids in order to build the bilayer (Fig. 5.2). These pairs were constructed with little or no interdigitation whilst achieving a near experimental value for the P-P interhead group distance of 34 +/ −1 Å.[2] These lipid pairs were then arranged onto both a hexagonal lattice and a cubic lattice (Fig 5.1 a-f) via rigid body translations. The placement of the lipid pairs was done by-hand and not an automated method, in order to attempt to reduce the regularity of the structure, which has been seen to reduce the accuracy of MD simulations in reproducing experimental results.

Table 5.1

Previous 5x5 Bilayer Average Temperatures (K)				
	Ave.	St. Dev.	Max.	Min.
PC	328.34	8.24	360.84	298.78
GL	289.82	10.18	330.31	258.77
Sn1	238.68	9.09	272.59	218.45
Sn2	239.47	9.53	281.12	221.59
Wtr	411.66	9.29	435.06	368.32
Sys.	311.91	10.92	353.43	272.60
Helix	315.48	3.03	324.45	307.15

5.1.2. Choosing the Bilayer Size - 6x6 or 5x5?

Molecular Dynamics simulations of lipid bilayers and monolayers have used various NxN lipid arrays, but N=3-5 models tend to cause artefacts. When periodic boundary conditions (PBC) are applied to such systems basic-ghost interactions across the unit cell are possible with a short cut-off, which can cause the system to have an induced crystallinity. As well as increased crystallinity, it is possible that when a solute is added to a low lipid number bilayer it may feel an effect from its ghost directly propagated through the solvating lipids. This effect reduces the ability of the

system to reproduce experimental data.

Using a cut-off for truncating the number of structural energy calculations requires that the solvent phase has a depth greater than the cut-off distance so as to disallow the interaction of the basic-ghost solute molecules through the PBC's. Therefore, a hydrated bilayer system requires a water layer of greater depth than the cut-off distance used.

To eliminate basic ghost solute interactions the cut-off distance can be reduced, which has the effect of speeding up each MD iteration. Again, as discussed above, this will have an effect on the efficacy of the simulation as the truncated interactions are important in stablising the structure.[3]

Increasing the bilayer size, from N=5 in a previous simulation, to N=6 in order to reduce the low lipid number affect, increases the total number of lipids to 22. This change dramatically increased the expense of the simulation because the number of waters required to solvate the head groups increased from 1028 to 1876. Thus, the increased number of lipids and water molecules caused a large reduction in the simulation speed with a bilayer of this size.

Instead of using a reduced cut-off to compensate for this decreased velocity of simulation a 16 Å cut-of was used so as to include longer range interaction to maximise the stabilising effect from the cross-bilayer head group interactions.

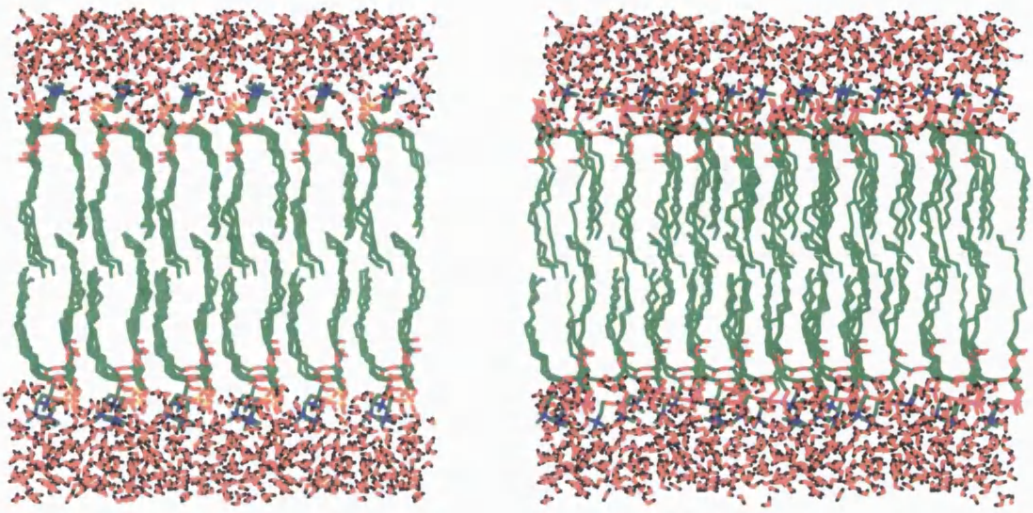
All these factors, whilst having a dramatic effect on the speed of the simulation, will help to produce a model DMPC bilayer that reproduces experimentally observed bilayer properties and can be used to study other bilayer systems.

Solvation of Bilayers

A 32 Å layer of water was used to hydrate the bilayers, so that when the unit cell was repeated in the periodic system, the multilamellar PBC arrangement was reduced, by the electrostatic cut-off, to an effectively uni-lamellar arrangement. The cut-off

disallowed water-head group interactions across bilayers. Head group hydration is important because the head group dipole moment extends ≈5 Å into the water phase and the charged head groups require a high number of associated water molecules to achieve full hydration. Such an inter-bilayer water layer satisfies the solvation requirement of the head groups which has a definite influence on the stability of the bilayer.[3] Thus, a 32 Å water phase combined with a 16 Å cut-off should reduce the chance of inter-lamellar interactions and increase the likelihood of a bulk water phase in the inter-bilayer aqueous layer that produce a stable head group region and bilayer structure.

The *Cell* option of *Insight* was used to assign new unit cell box dimensions to the hexagonal and cubic anhydrous bilayers, with orthorhombic angles (α, β, γ = 90°) and the *Soak* command was used to place pre-equilibrated water molecules in the defined box according to their Van der Waal's (VdW) radii. The box dimensions used were those of the anhydrous bilayer with 16 Å added to the maximum and 16 Å subtracted from the minimum y-coordinate(the bilayer normal). The x and z-coordinates, which describe the bilayer plane, were maintained from the anhydrous bilayer. The water molecules placed in the interior, hydrophobic region of the bilayer were removed up to the C_6 of the glycerol moiety as it has been shown that the head group waters penetrate up to the carbonyl oxygens of the glycerol moiety[4,5]



a b

Fig. 5.3 - Initial Unit Cell Coordinates, Sq (a) Hex (b)

High energy, clashing waters were present at the edge of the hydrated box due to inaccuracies in the Soak procedure which measures the density of the box and then uses the VdW radius of water to place molecules in the described area up to the required density of 1 gml^{-1}. Thus, at the unit cell boundary waters can clash after the box is placed into a periodic system. Careful use of the minimiser released most of the energy induced by the boundary clashes and only very high energy waters remained. This small number of high energy waters were removed manually via *Insight*, which had a negligible affect on the density of the aqueous phase. During this process the bilayer was held fixed, allowing only the waters to minimise. The full unit cell, including the hydration layer, is illustrated in Fig. 5.3.

5.1.3. Minimisation

The two bilayers were minimised to reduce the strains induced in the general building process and to remove clashing waters, as described above. The minimisation was done using the steepest decents method for a short period to remove the low level strains that resulted from the building process. The lipids have been constructed using lipid acyl chains taken from a dynamic system and therefore, minimising them to low derivative values could remove the important dynamic characteristics they possess.

The systems were minimised for 100 and 150 iteration respectively, using the PBC dimensions modified from the soak procedure.

5.1.4. Dynamic Sampling

The molecular dynamics trajectory of these two systems were sampled using the V.F.F., Valence Force Field, with a time step of 1 fs. As described before, a truncation cut-off was used for the non-bond interactions and the cut-off distance was set at 16 Å. An 18 Å distance was used as the cut-off for the neighbour list generation for the non-bond calculation.

The unit cell vectors (UCV) used were maintained from those used in the minimisation process. The UCV dimensions were then varied according to the refinement process explained below. A multiple temperature bath was used during all the described trajectories with the temperature varied according to the requirements of the simulation and with one temperature bath being applied to the bilayer atoms and a second to the water molecules. All the sampled trajectories were run at a temperature bath value of 311 K.

5.1.5. Bilayer Refinement Protocol

After the building and the minimisation of the systems, a refinement process was used to improve the reproduction of experimentally observed data.

In order to attempt to achieve the experimental surface area per lipid head group (SA) and the cross bilayer, inter-head group phosphorous to phosphorous (P-P) distance, a series of adjustments were made to the system volume and therefore, to the effective SA. The P-P distance was monitored in order to investigate its coupling to the SA.

Table 5.2

Original structural values		
Sys.	SA/lipid ($Å^2$)	P-P distance (Å)
Hex	86.4	37.6
Sq	75.3	37.6
Expt.	*60-65*	*34+/-1*

It was postulated that reducing the y-axis dimensions (the lipid long axis - bilayer normal) will reduce the P-P distance as a consequence. The SA change could also alter the P-P distance as an increase/decrease in the available space will allow the acyl chains to increase/decrease their rate of torsional isomerisation. Thus, as the isomerisation rate maybe correlated to the P-P distance, a change in the SA should effect a change in the P-P distance.

The two original bilayers have the SA and P-P values shown in Table 5.2 above. The close packing of the cubic system can be seen and as well as the larger head group area of the hexagonal geometry. This basic geometrical data might suggest that the close packed lipids of the cubic based system may induce more intermolecular collisions, as the count of collisions is correlated to the density of the system. It has been shown that in a decane bilayer system that the density of the decanes is inversely related to the fluidity of the molecules,[6] thus, as the lipid bilayer interior is a milieu of

decane like chains, this is in agreement with the above postulate on acyl chain dynamics.

Unit Cell Refinement

During the unit cell refinement process, NPT conditions with a multiple temperature bath applied were employed, which allowed the UCV's to be scaled. UCV scaling was used to change the unit cell volume and therefore, the SA. The UCV's were gradually scaled over 25 ps, with the temperature gradually increased from 200 K to the simulation temperature of 311 K over the same period. This temperature incrementing was carried out in order to reduce high energy, potentially artefactual clashes that may result from the increasing density.

The UCV of the Hex system were reduced in both the x and z-axes by 7.32 Å over the first 20 ps and the y-axis independently by 1.85 Å. The UCV of the Sq system were reduced in both the x and z-axes by 2.2 Å and in the y-axis by 1.7 Å over the same 20ps. The SA and P-P distance values were monitored throughout this period of the simulation in order to map how the changes in the unit cell volume and then the free NPT conditions affected them.

After 20 ps the Hex system had reached a reasonable structure with a SA of 64.83$Å^2$ and a P-P distance of 35.48 Å, therefore both are near the experimental values. At this point in the simulation the compression of the system caused a disproportional increase in the system pressure for a small change in the UCV, this suggests that this volume is near the correct value for this system under the V.F.F.. The Cubic system after the initial 20 ps scaling dynamics had achieved a reasonable SA value of 62.32 Å, whilst the P-P distance was 37.8 Å, thus it remained near the original value.

The energies of the two systems at 20 ps indicates that the cubic system is less stable, as it has a total energy of -13962 kcal whereas the Hex system has a total energy of -17136 kcal. Thus, combining the energetic data to the structural data the

Hex system has a better structure than the Sq system which has a higher than required P-P distance, as well as a higher energy.

It was decided to investigate further the relationship between the SA and the P-P distance by increasing the Sq SA to the Hex bilayer value at 20 ps and then reducing the Hex SA to the Sq system value at 20 ps. This procedure was undertaken in order to confirm the relationship observed above, where the Hex bilayer SA has a sensitivity to reducing volume, whereas, the Sq bilayer has no sensitivity.

From 20 ps the Hex system had its x and z-axis scaled down by 2.5 Å and the Sq system had its x and z-axis scaled up by 2.5 Å. This scaling was done over a 10 ps period under NPT conditions and at 311 K and was followed by another 10 ps of simulation under NPT conditions without volume scaling, with the UCV's allowed to fluctuate. Thus, the reaction of the system to the second period of scaling was followed.

The scaling up of the Sq system resulted in a P-P distance of 38 Å, a negligible change, whereas the Hex system scaling produced a reduced P-P distance of 34.5 Å. After the scaling, both the systems showed that they were not stable as they gradually relaxed back to near their SA values at 20 ps. The P-P distance for the Hex system also relaxes back to near its original value, but the cubic system shows only a minor change to 37.7 Å. Thus, it was concluded that changes in unit cell volume effect only the less dense, Hex system and that its SA is related to its P-P distance. Interestingly, the more dense cubic system shows little correlation between its SA and P-P distance. The results for the Hex system agree with the postulate that density is inversely related to acyl chain dynamics, but the results for the Sq system appear to contradict the postulate. This could be due to an event occurring which prevents the lipids from fluidising their acyl chains.

Following the refinement procedures, the bilayer structures at 20 ps still best reproduced the experimentally observed data therefore, the coordinates at this point were used for the simulations that follow.

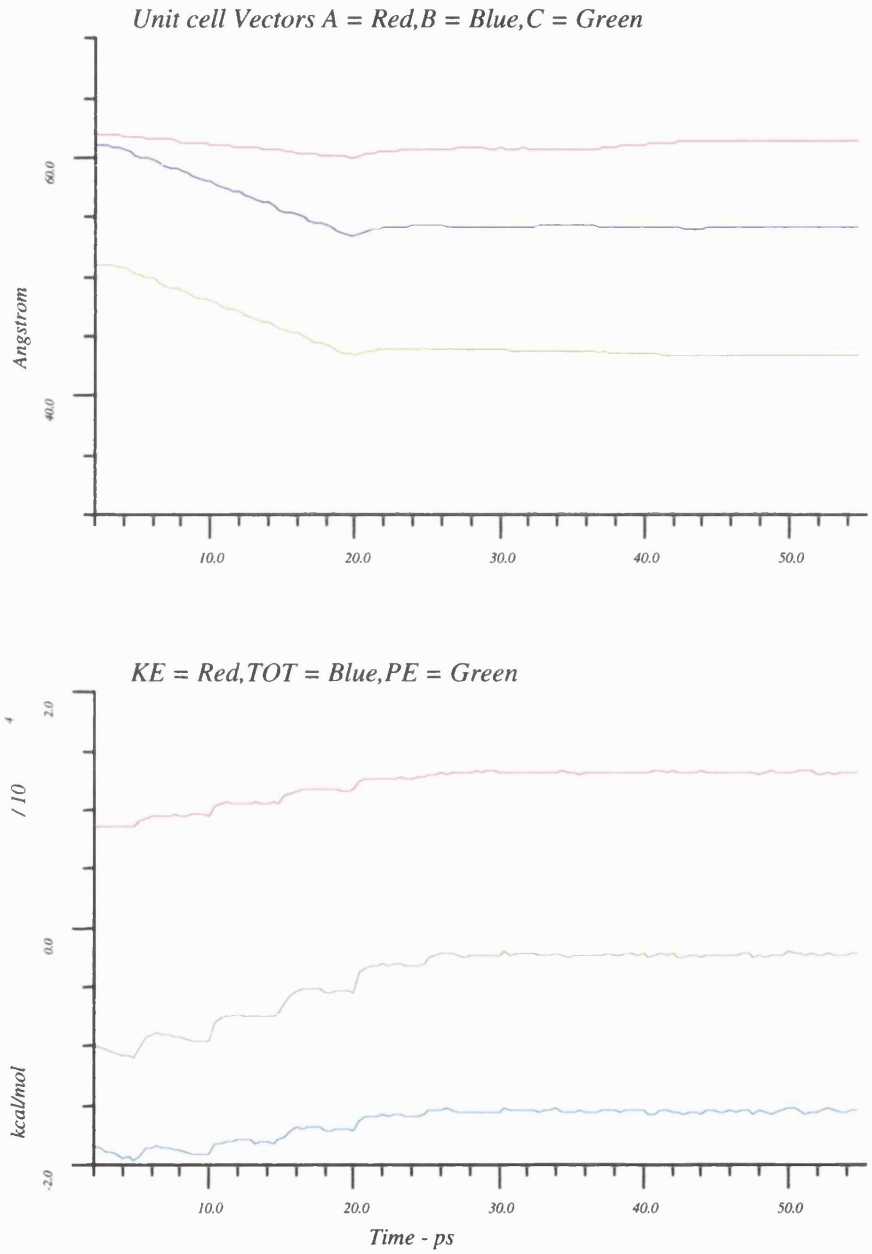


Fig. 5.4a Hex - Thermodynamic Results

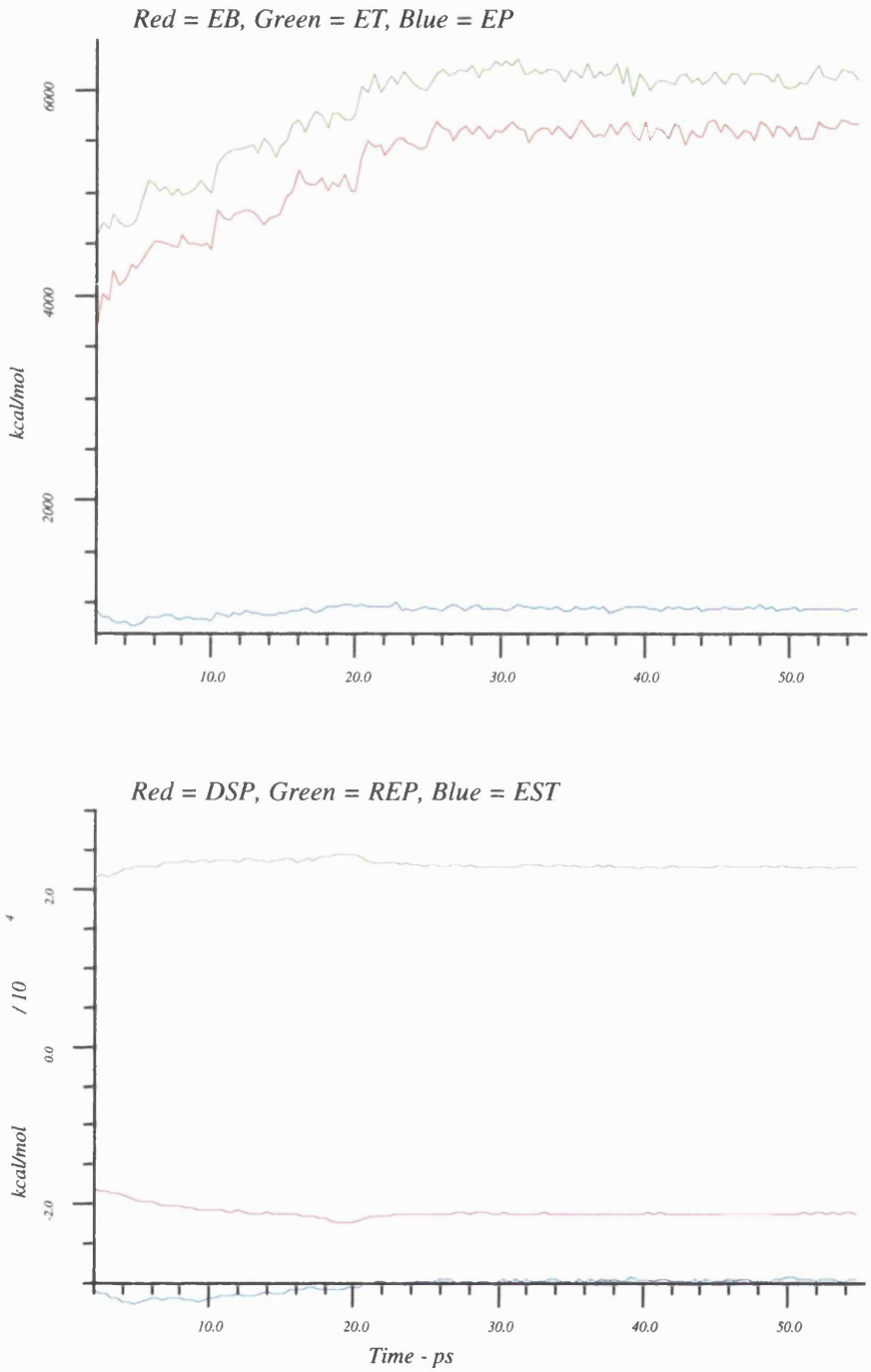


Fig. 5.4b Hex - Thermodynamic Results

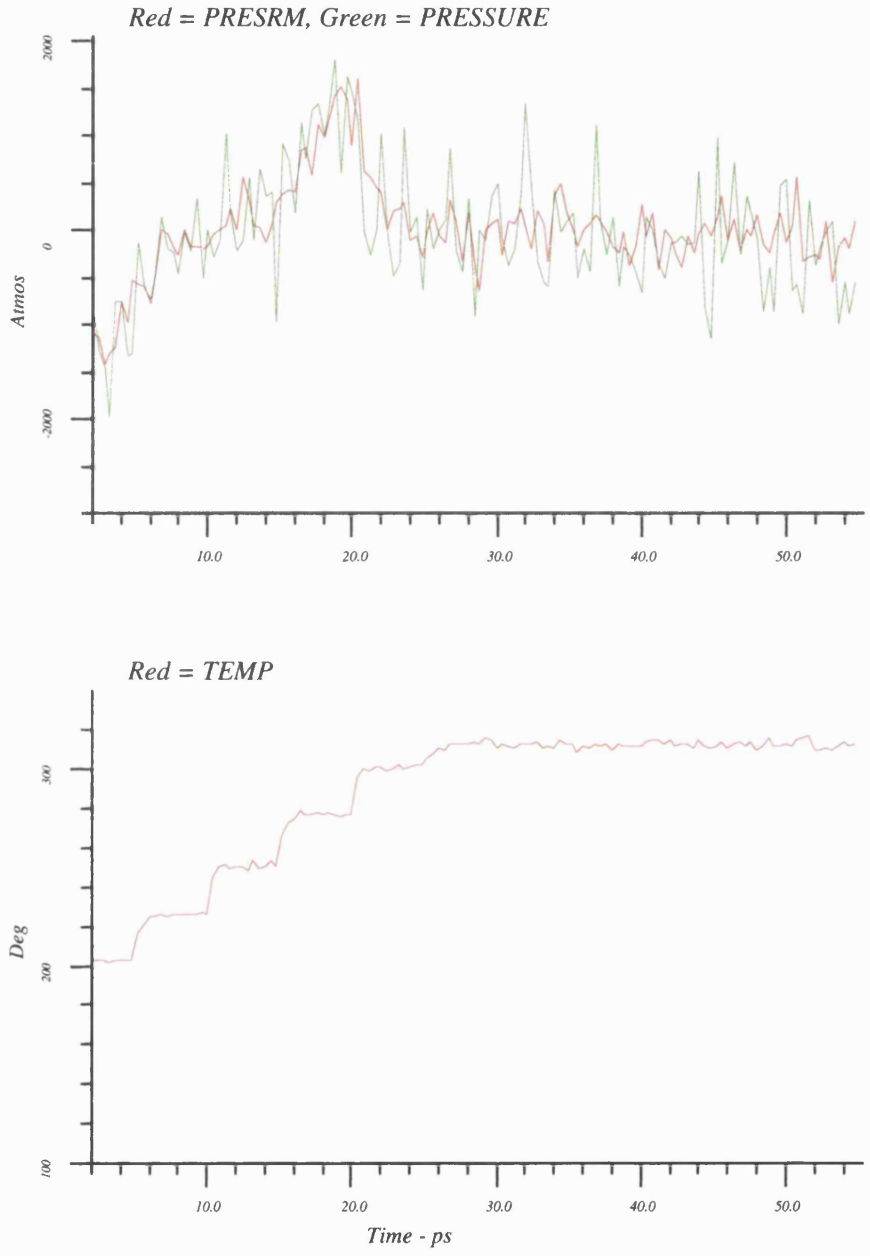


Fig. 5.4c Hex - Thermodynamic Results

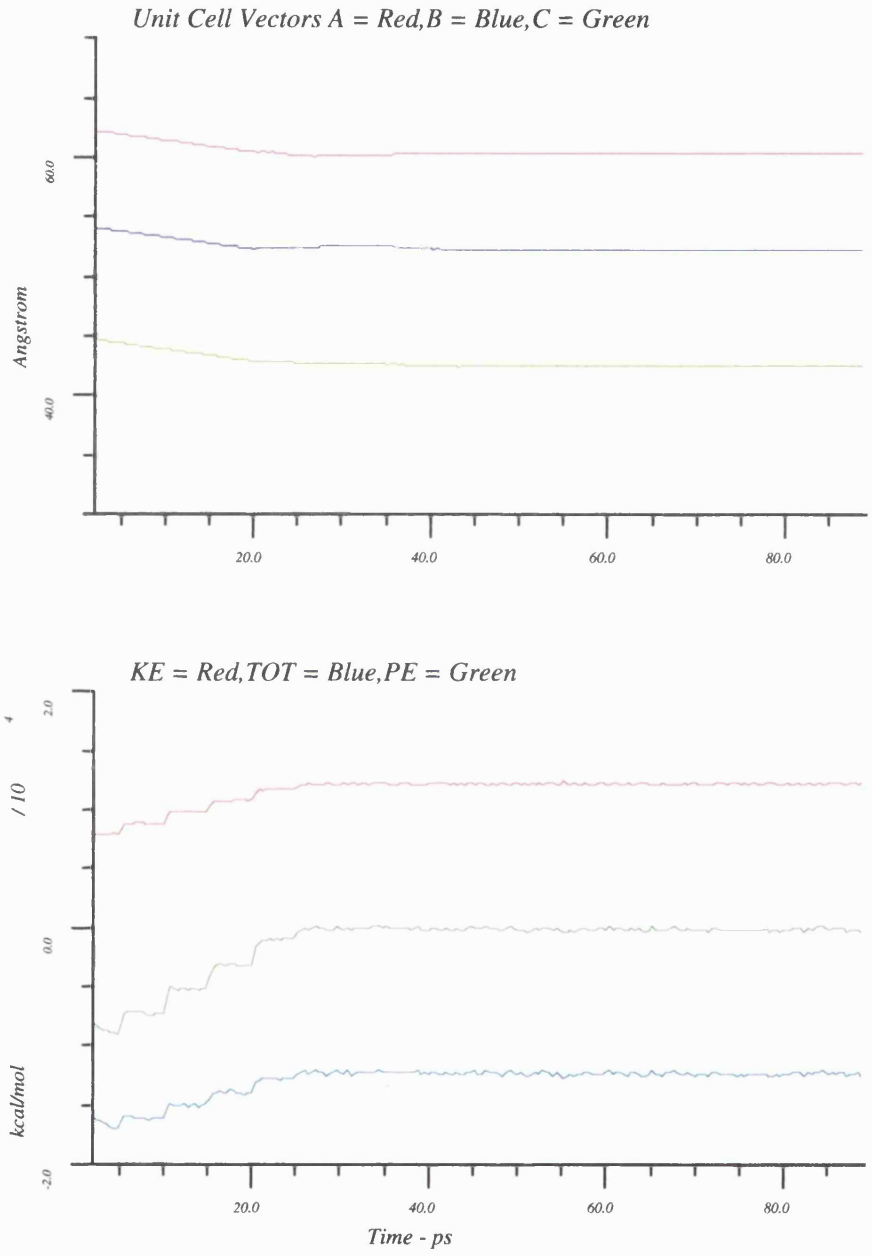


Fig. 5.5a Sq - Thermodynamic Results

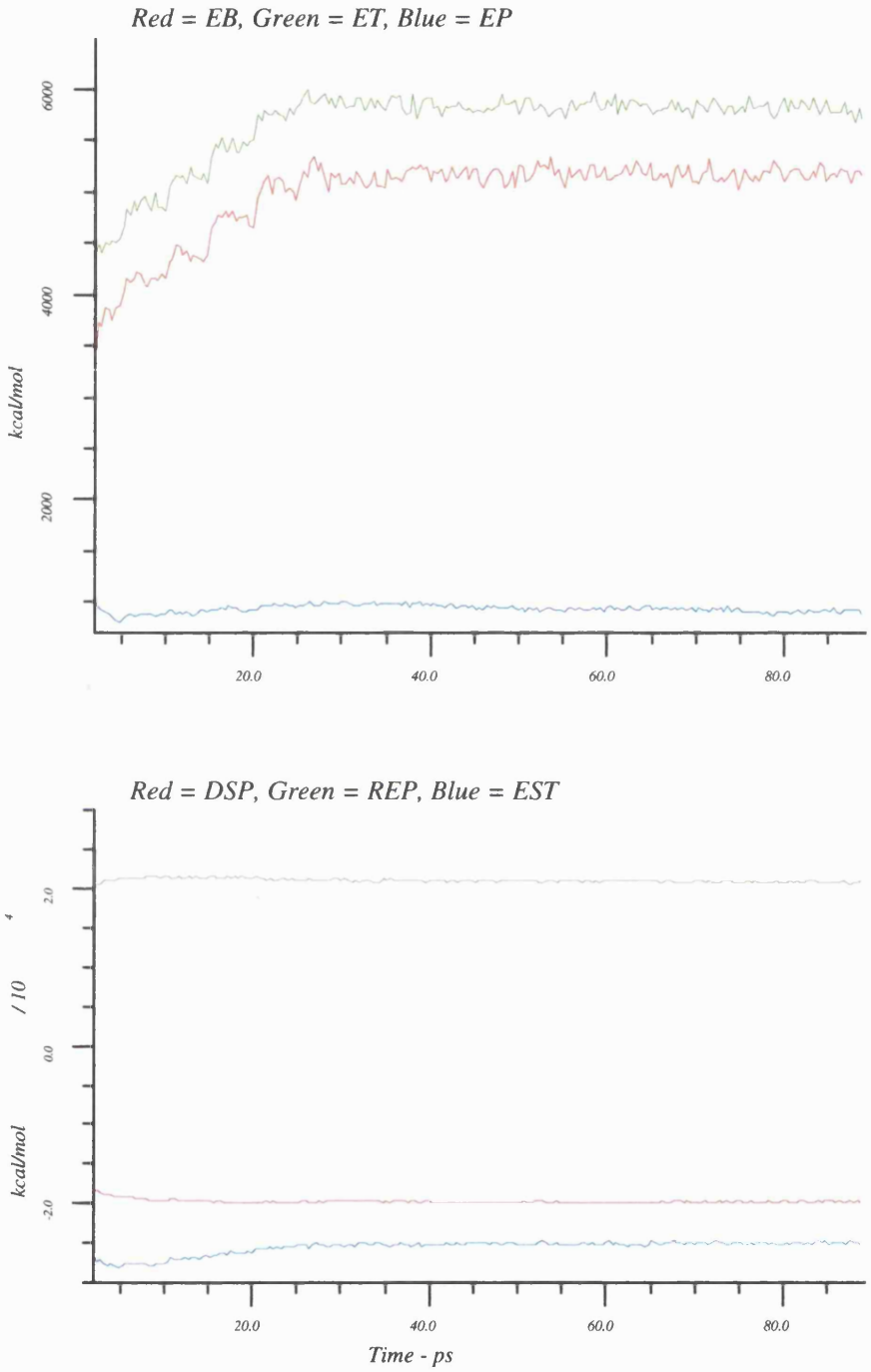


Fig. 5.5b Sq - Thermodynamic Results

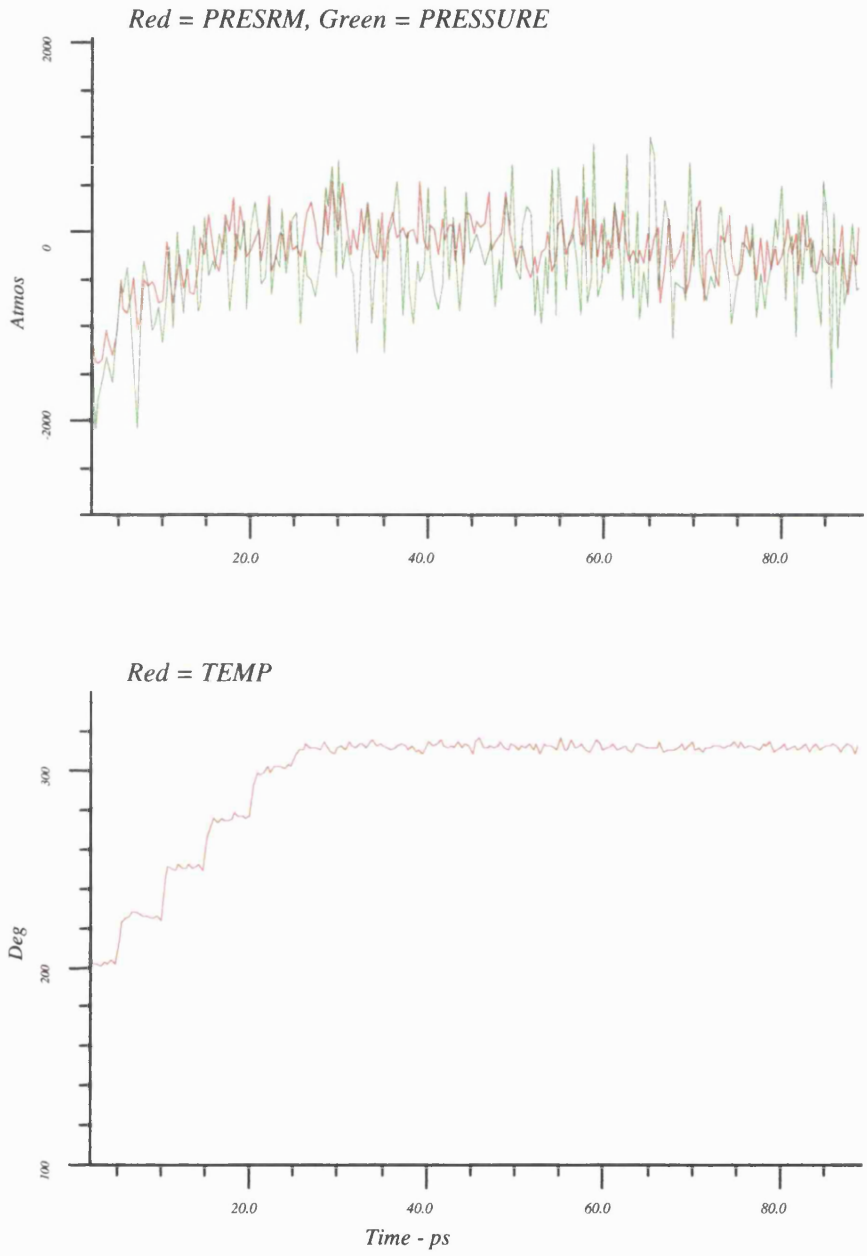


Fig. 5.5c Sq - Thermodynamic Results

In order to further investigate the stability of the two structures they were run for 15 ps under NPT conditions, with a constant pressure of 1 atmosphere and a multiple temperature bath of 311 K. The thermodynamic plots (Fig. 5.4 and Fig. 5.5) show the UCV remain reasonable stable during this period indicating that the SA are correct for the two systems under the parameters of the V.F.F. force field. The UCV's of the Sq systems remained quite constant whilst the Hex UCV's increased slightly during the initial period then remained essentially constant.

The V.F.F. programme uses a pressure bath which couples pressure changes to the unit cell vectors. Even though the average pressure is around 1 atmosphere the instantaneous pressure, with or without the pressure bath, ranges from -2000 to +2000 atmospheres. The coupling of these large changes to the UCV can cause extremes in local pressure and may cause artefactual changes in the systems. We used the NPT conditions in an attempt to reduce the volume of the system.

Taking into consideration the problem of maintaining the favourable physical data, such as SA and P-P distances, with the pressure bath induced fluctuations in the UCV it was decided to use a constant volume ensemble instead of constant pressure. This ensemble would preserve the desired physical box dimensions whilst removing the possibility of pressure bath induced artefacts. The plot of pressure, before and after the change from NPT to NVT conditions shows, little change and illustrates that even though the average of the NPT section is 1 atmosphere the observed fluctuations of $\pm$1000 atmospheres makes this a less meaningful value.

Thus, the 2 systems were continued under the NVT conditions and at 311 K for 220 ps for the Hex system and for 90 ps for the Sq system.

5.2. Results

All time average results were calculated for the Hex simulation between 50 and 220 ps and for the Sq simulation between 50 ps and 90 ps.

5.2.1. Hexagonal Packed Bilayer - Hex

5.2.1.1. Average Temperature

The temperature plots for the atoms comprising the different sub-sections of the bilayer system are illustrated in Fig 5.6.

Fig. 5.6 shows that the temperature of the different groups, are close to the temperature bath value, but Table 5.3 show that the averages are slightly above or below their required value, with the system average being slightly above 311 K.

The unit cell atoms can be divided into two sub-groups: (1) non-polar - the acyl chain atoms and (2) polar or charged - the Choline, Glycerol and Water atoms. Cut-off methods cause an artefactual temperature change according to the polarity of the atoms involved. The groups containing polar atoms are warmed by the inconsistencies in the structural energy calculation, the cut-off induced temperature effect, giving an average value above that expected and conversely, the groups containing non polar atoms tend to be cooled by the cut-off induced temperature effect giving an average value below that expected.

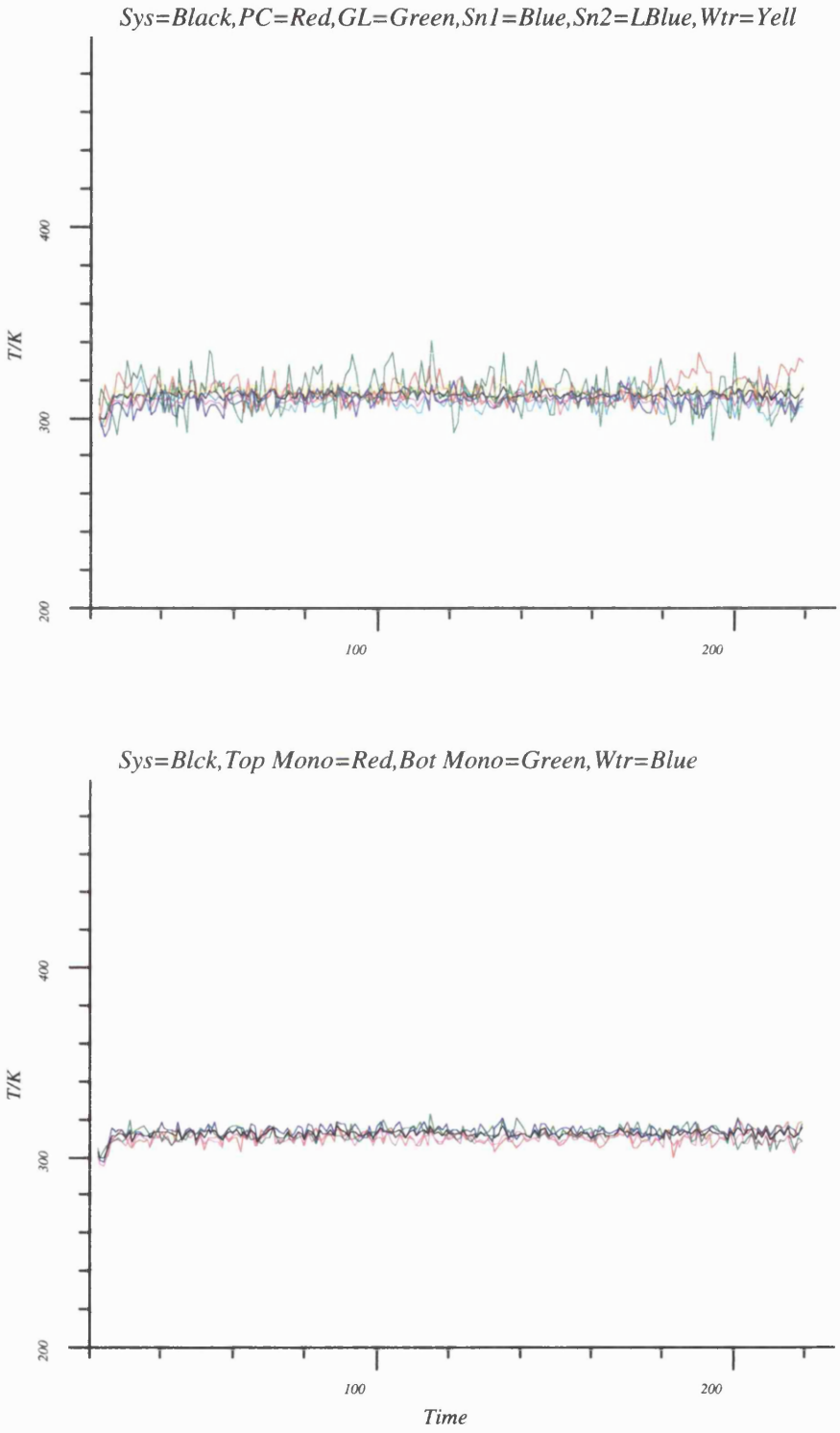


Fig. 5.6 Hex - Average Temperatures

Hence the highly polar atoms will have a higher than average temperature and the apolar atoms a lower than average value (Table 5.1) even though the average temperature is essentially that required. The multiple temperature bath improves the efficiency of this method, but still leaves the temperature 2-3 degrees from that required which is inevitable if each polarity group does not have its own temperature bath. Thus, 5 or 6 baths maybe needed, otherwise such heterogeneous systems will have temperature differentials between different types of atoms. The temperature of the system, 312 K, even though it is not quite at the required 311 K temperature, is still is within the L_α range of the bilayer phase diagram.

Table 5.3

Hex, Average Temperatures (K)		
Group	Temperature	S.D.
System	312.1	2.1
Choline	314.0	6.2
Glycerol	313.6	9.2
Sn1	309.6	4.6
Sn2	309.3	4.5
Water	313.7	2.9
Acyl Chain	309.5	3.0

5.2.1.2. Thermodynamic Data

The thermodynamic data is shown in Fig. 5.7, with the first 50 ps included separately in Fig. 5.4.

The thermodynamic data shows that the NPT section of the simulation from 20-30 ps reduces the slight strain induced during the unit cell vector (UCV) scaling procedures. The pressure plot (Fig. 5.4) clearly shows that the pressure settled, from its peak at 20 ps, to a reasonable value around 1 atmosphere, as a slight increase in the UCV's relaxes the pressure. The UCV's stay stable after this readjustment, indicating that there are no appreciable change to the thermodynamics when the NPT conditions are changed to NVT conditions.

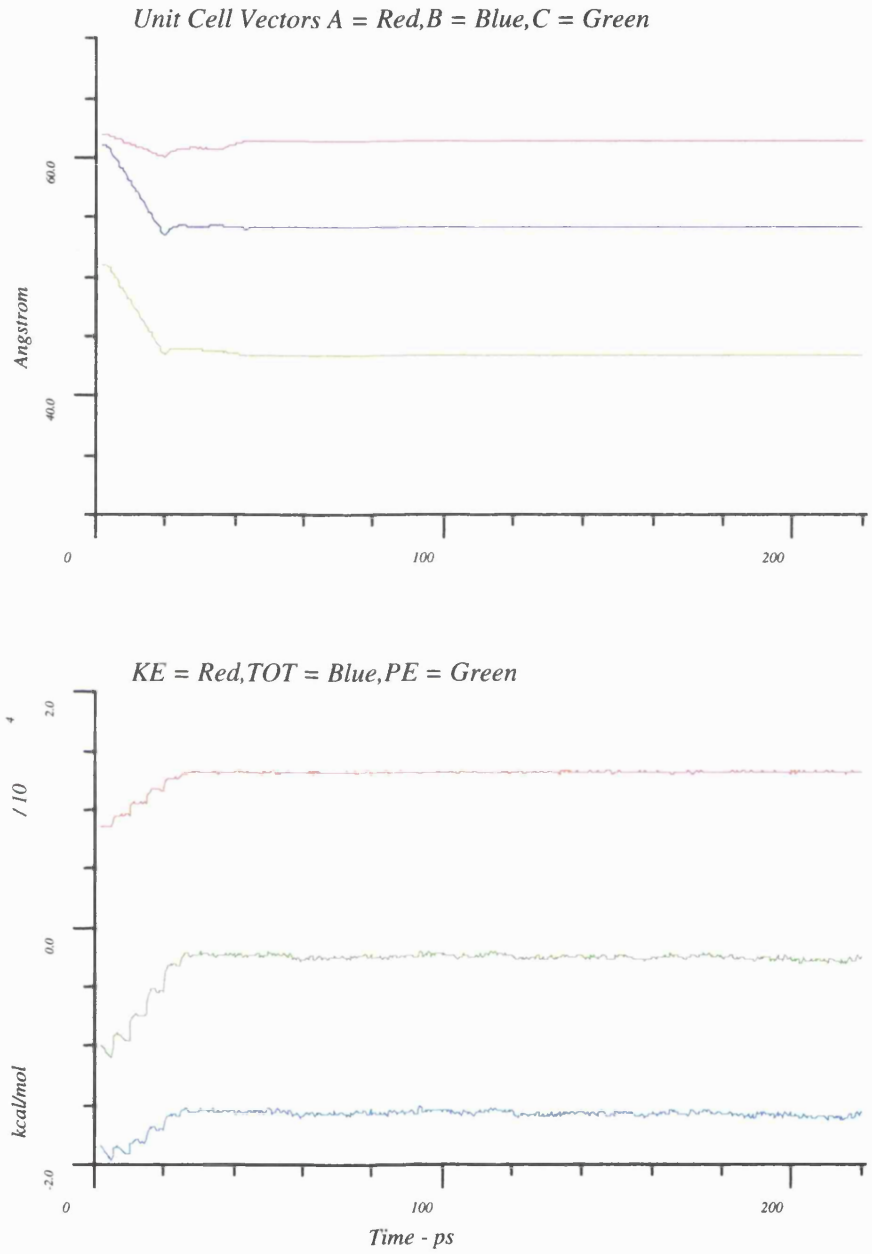


Fig. 5.7a Hex - Full Thermodynamic Results

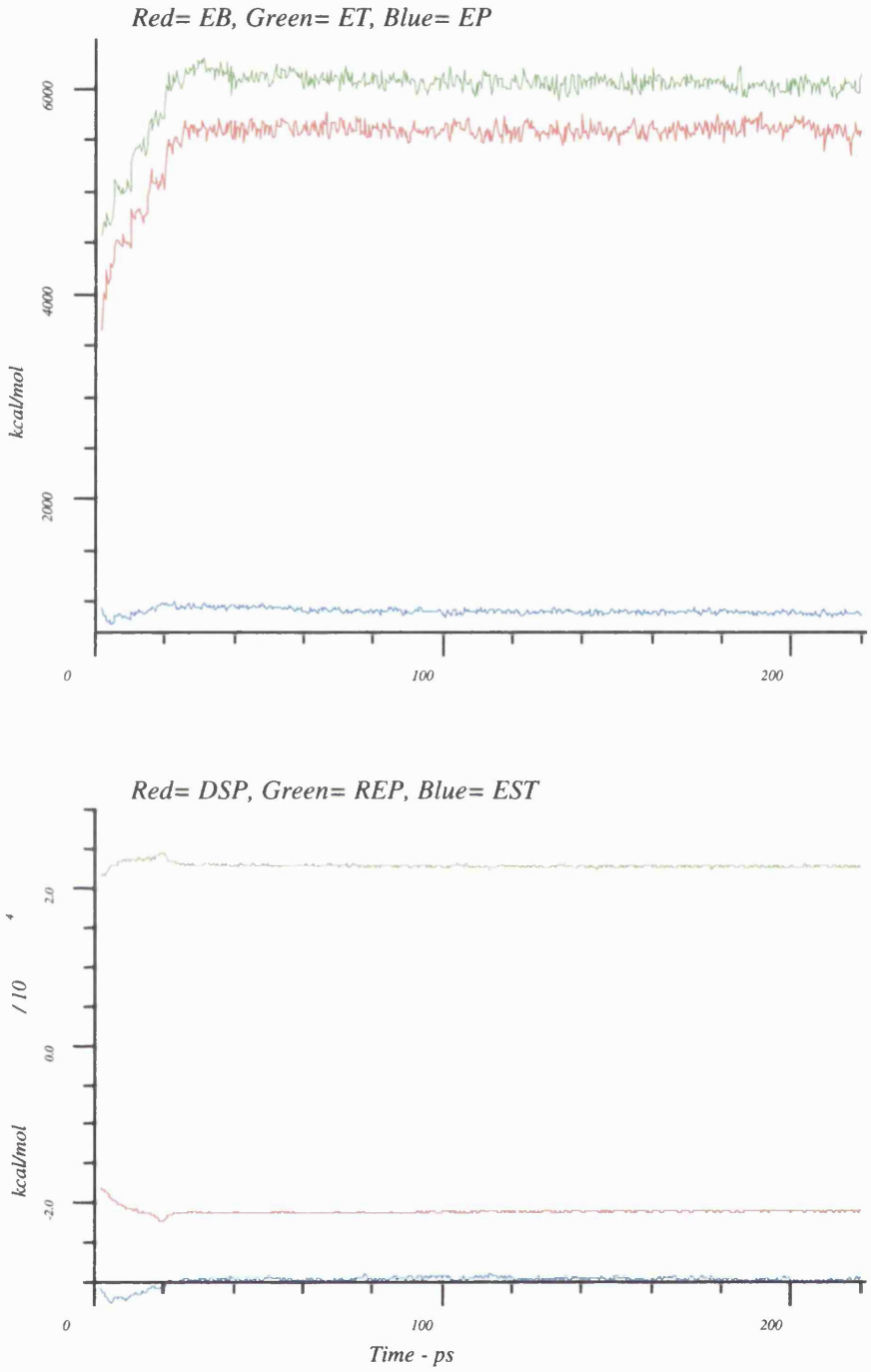


Fig. 5.7b Hex - Full Thermodynamic Results

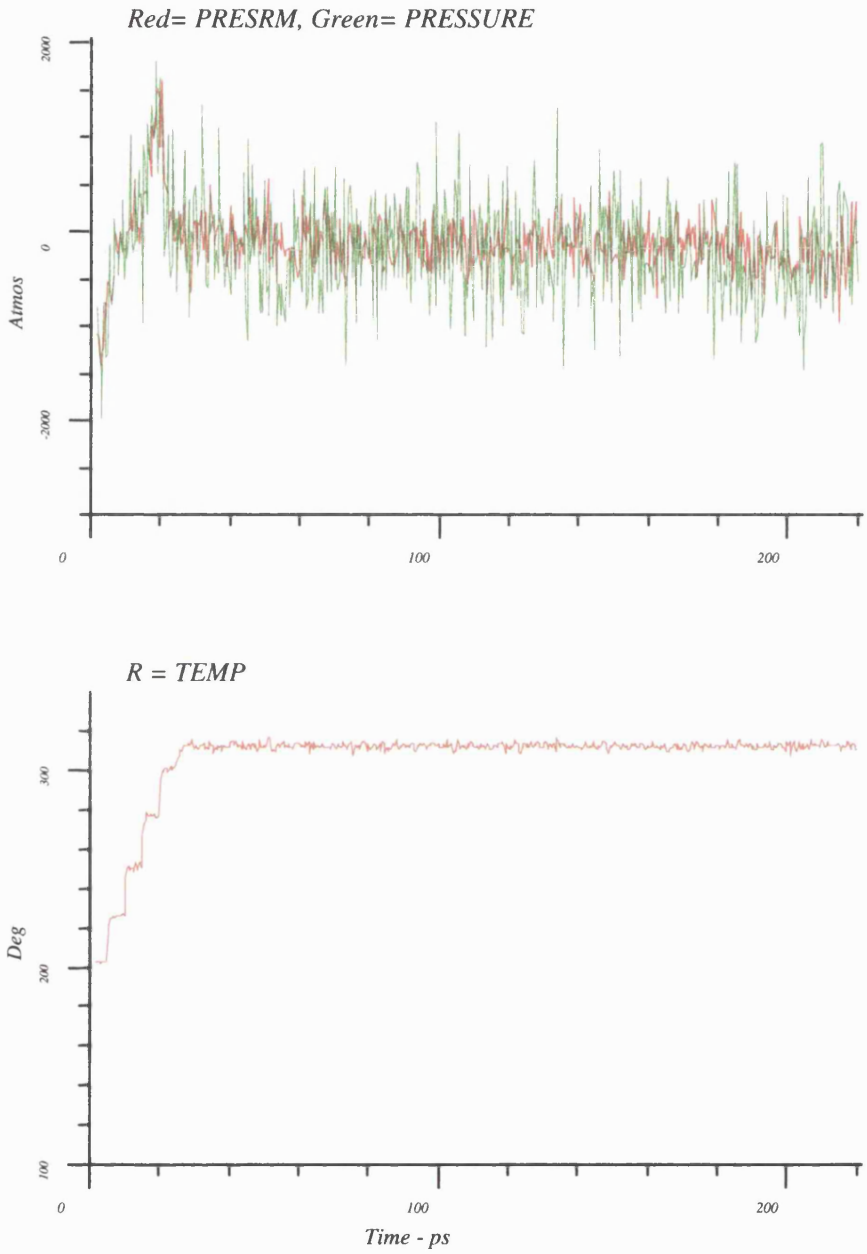


Fig. 5.7c Hex - Full Thermodynamic Results

The energetics are reasonably stable throughout the simulation and only show variation due to the warm-up phase in the initial stage of the simulation. The non-bond energies show changes associated with the reduction in the volume caused by the UCV scaling procedure. The repulsive term shows a slight rise as the unit cell density drops and the free volume available to the atoms reduces. A concomitant change in the electrostatic energy occurs probably as a result of the head group structure becoming a little less crystalline. The dispersive term of Lennard-Jones potential shows a more marked change during the UCV scaling procedure. The increasing density obviously favours the attractive Van der Waal forces and is a contributor to the stabilising of the lower SA per lipid structure.

The thermodynamic data therefore indicates that the system is an equilibrated state and undergoing no major structural changes.

5.2.1.3. Density Profile Function Data

The density profiles (DPF) for the sub sections of the bilayer are illustrated in Fig. 5.8 for the Hex system.

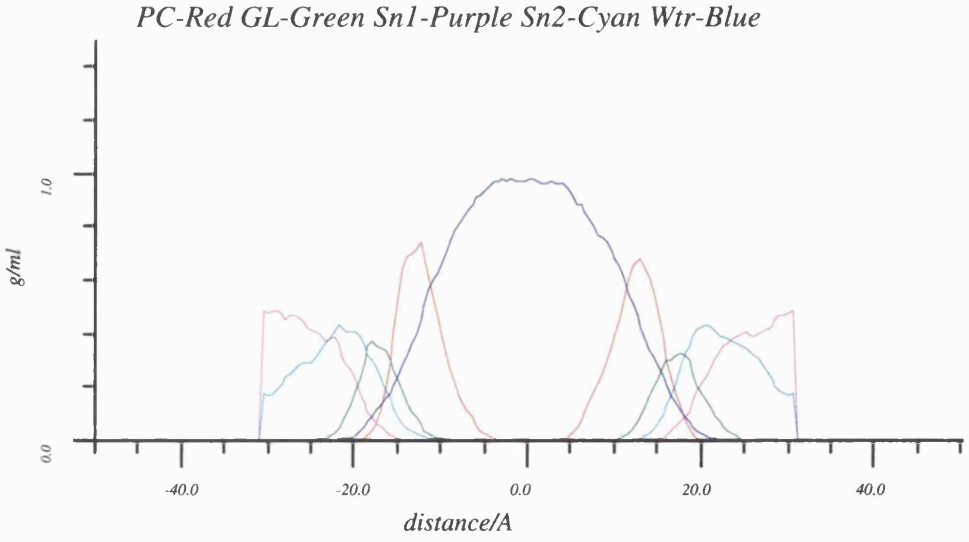


Fig. 5.8 Hex - Density Profile

The profile clearly indicates that the inter-bilayer water has a 10 Å wide layer of bulk structure with a density of 1 g mol^{-1}.

The typical rough bilayer interface can be seen to be present in the head group region (±20 Å to ±5 Å), as is indicated by the variation in head group penetration into the water layer and the significant probability of finding water molecules in the glycerol region of the bilayer. The two monolayer choline densities indicate that the most probable displacement of the inter-bilayer head groups is at a 25 Å separation. Thus, the head groups maintain an average displacement above the cut-off distance.

Normally the bilayer is divided into 2 main groups - the acyl chains and the head group region, but according to the Marrink and Berendsen classification method, the bilayer has four main sections[7] which divides the bilayer more accurately into sub-sets according to density and their physical properties:

Section 1 - the wide and diffuse (rough) bilayer surface, where the bilayer disrupts the waters structure, is clearly observed.

Section 2 - spans the area where the head group density is high, with a width of 0.8 Å, similar to the authors value of 0.75 Å.

Section 3 - possess the high acyl chain density, an area which spans the region from the end of the penetrating waters to the bilayer interior, where the density drops to that of hexadecane and has a width of 0.7 Å, which is in agreement with the authors.

Section 4 - the low density region, where the acyl tail density drops to that of liquid hexane and we get a bilayer width of 1.2 Å, only a little above that of the authors.

Thus, the Hex bilayer has a good structure and is in agreement with published

data.

5.2.1.4. Lipid Structural Internals

Table 5.4 shows the average values and standard deviation (S.D) for the selected structural angles that describe the internal motion of the lipid molecules in the Hex simulation.

Table 5.4

Hex, Lipid Structural Angles		
Number	Value	S.D.
θ_1	109.23	12.62
θ_2	46.66	25.35
θ_3	103.66	30.12
θ_4	126.53	23.49
θ_5	108.97	24.24
θ_6	128.20	19.24

The S.D values indicate that the lipids have mixed flexibility, indicating a good degree of heterogeneity in the bilayer.

θ_1 shows that the head group is perpendicular to the bilayer normal but directed slightly into the water phase.

θ_2, indicates that there is considerable acyl chain motion relative to the head group which could be caused by a scissoring action of the two acyl chains or by torsional isomerisations in the acyl chains. The value for θ_2 when the acyl chains are aligned and all trans is $\approx$ 20-30 °.

The data shows that the bend in the Sn2 chain, imparted in order to bring the acyl chains in line, is maintained through the simulation. This is seen in $\theta_{3,5}$ (associated with the Sn2 chain) which are less than the $\theta_{4,6}$ (associated with the Sn1 chain). The S.D. values also indicate that there is a difference in the mobility between the Sn1 and Sn2 chains, which again could be an effect of the different positioning of the acyl chains on the glycerol moiety.

Table 5.5 shows the average vales and S.D for the selected structural lengths that describe the internal motion of the lipid molecules in the Hex simulation.

Table 5.5

Hex, Lipid Structural Lengths		
Number	Value	S.D.
d_1	4.74	0.20
d_2	6.94	0.64
d_3	17.21	1.47
d_4	15.26	1.90
d_5	14.03	0.80
d_6	13.72	1.03
d_7	12.03	4.64
d_8	36.02	2.65

The S.D values indicate that lengths 1,2,5 ($d_{1,2,5}$) have a reduced motion over the lengths 3,4,6,7,8 ($d_{3,4,6,7,8}$). $d_{1,2}$ describe the motion of the head groups and are in agreement with the accepted semi-crystalline structure of DMPC head groups.[8]

$d_{3,4}$ have dissimilar values and S.D. indicating a slightly higher flexibility in the Sn2 chain over the Sn1 chain. $d_{5,6}$ have similar values but indicate slightly higher motion and mobility in the Sn2 over the Sn1 chain. This slightly higher Sn2 motion is in general agreement with the conclusions of the structural angle data, but the difference is more obvious in the angle data than in the above length data. The difference in the mobility indicated maybe a result of flexibility in the glycerol region as $\theta_{3,4}$ have a contribution from the glycerol moiety and are more flexible than $\theta_{5,6}$ which do not include the glycerol atoms. The acyl chain $d_{3,4}$ S.D. values, which include the C6 glycerol atom, have a higher mobility than the $d_{5,6}$, which only include methylene and methyl atoms. Clearly, these results indicate that the glycerol confers a degree of flexibility on the lipid molecules.

d_7 clearly indicates a very flexible inter-acyl chain termini length and therefore a high degree of activity in this mode of lipid motion. This is either due to the trans-gauche isomerisation of the acyl chain torsions causing them to shorten or lengthen or

the two acyl chains could be undergoing a scissoring motion. The structural data indicates that either of the two modes are possible as trans-gauche flips could add up to such a displacement of the acyl chain ends, but the flexibility in the glycerol region could equally indicate a scissoring action.

The P-P distance described by the d_8 length shows a value of 36 Å and has a fairly high S.D.. This value is a little above the experimental value of 34 +/-1 Å, but this value agrees well with other MD simulations.[9]

5.2.1.5. Segmental Order Parameters

Fig. 5.9 shows the segmental order parameters (S.O.P.) for the acyl chain methylene groups of the Hex system.

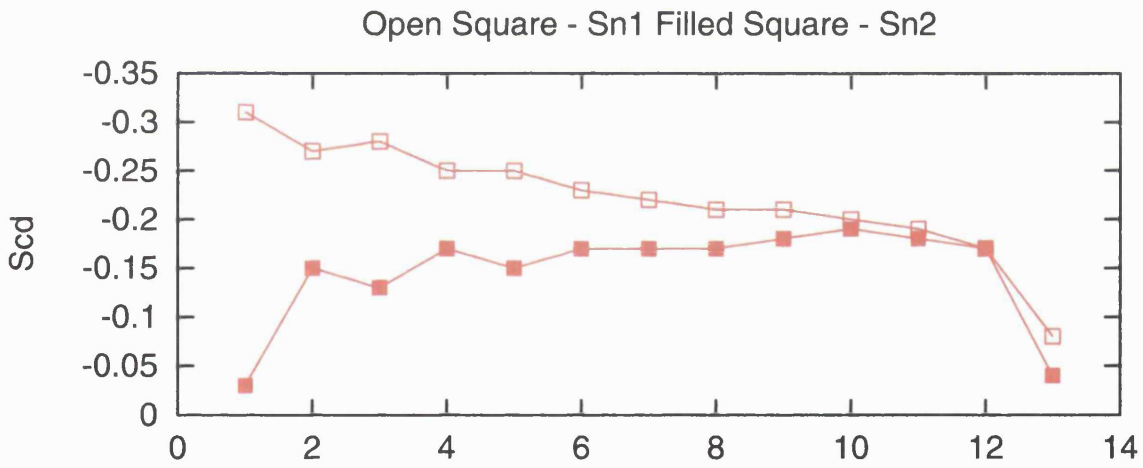


Fig 5.9 Hex - Segmental Order Parameters

The most obvious result indicated by the plots is the difference in dynamic motion between the Sn1 and Sn2 chains. The higher order parameters for the Sn1 chain agrees with the structural data, indicating reduced motion in the Sn1 methylene groups which is concentrated around the upper region of the chain.

In general the S.O.P. profile shows good agreement with experiment,[2] with the main characteristics of the experimental data being replicated in the results:

i. The lower order parameter of the C22 carbon.

ii. The order parameter plateau through the middle section of the chains.

iii. The high flexibility of the acyl chain terminal methylenes.

The first characteristic arises from the geometry of physical L_α phase lipids, which is maintained from its crystalline structure, with the Sn2 chains being bent at its first chain position in order to produce the parallel alignment of the acyl chains. The bent structure causes the methylene group to be non-aligned with the bilayer normal, therefore giving the upper Sn2 methylene an artificially low order parameter value, but the order of the Sn1 upper chain region is governed by a balance of intra and inter-molecular interaction which disallow low order (explained below).

The second and third characteristics arise from a combination of inter and intramolecular effects. The population of trans gauche (t/g) conformers along a chain should be uniform as its energy barrier is standard at 2.1 kj mol^{-1} per methylene,[10] independent of the chain position. This intramolecular property is affected by the intermolecular consequence of such torsional changes. In the upper part of the acyl chains, when a trans-gauche torsional flip occurs the chain will move a long way relative to the head group entity (essentially fixed due to its large attraction to water). Thus the moving chain will have a high probability of clashing with an adjacent acyl chain, substantially reducing the chances of this conformational change occurring. Nearer the ends of the acyl chains such a torsional flip has a low probability of causing

such a clash and therefore the likelihood of a t/g isomerisation is at a maximum at the terminal methylene.

5.2.1.6. Torsional Data

Table 5.6 contains the trans-gauche torsional population (t/g) data for the methylene groups of the Sn1 and Sn2 acyl chains.

Table 5.6

Hex, Lipid Acly Chain t/g Percentage							
Sn1				Sn2			
Torsion No.	t	g+	g-	Torsion No.	t	g+	g-
14	83.65	6.70	6.60	29	37.84	4.91	51.94
15	80.93	5.25	10.67	30	72.03	2.09	22.42
16	83.74	4.91	8.70	31	76.87	4.33	15.88
17	85.64	3.74	8.19	32	74.59	3.92	18.45
18	88.53	3.15	5.92	33	83.92	1.54	11.95
19	83.21	5.22	9.09	34	82.30	1.33	13.87
20	89.05	4.41	4.14	35	85.54	3.63	8.68
21	86.86	5.41	5.47	36	85.96	4.11	7.56
22	89.48	3.86	4.51	37	82.55	6.16	8.88
23	85.83	4.11	7.64	38	88.52	5.36	3.96
24	87.06	4.18	6.42	39	88.39	4.28	5.08
Ave.	85.82	4.63	7.03	Ave.	78.05	3.79	15.33

The higher flexibility in the Sn2 upper chain over the Sn1 chain can be observed in Table 5.6, with torsion 29, which corresponds to the bend in the Sn2 chain, being highly gauche. This behaviour is also observed in torsion angles 30-32 of the 2nd, 3rd and 4th methylene groups, which have significant gauche populations and again, illustrate the low order seen in experiment. [2] The Sn1 t/g values are again a little below the Sn2 values, this is in agreement with the SOP and the structural internal results.

The original acyl chain, i.e. those extracted from the previous simulation, contained 2.4 gauche conformers per chain, in good agreement with experimental value of ≈ 2.5[11] to 3. [8] The Hex simulation has an average value of 1.3 and 2.3 in the Sn1 and Sn2 chains respectively. Thus, both these values are below the expected value. The

possible reasons for this reduction are either due to the building process, the refinement procedure or the main simulation and will be discussed later in the Chapter.

Table 5.7 contains the standard deviations (S.D.) data for the lipid torsion angles of the Hex simulation.

Table 5.7

Hex, Lipid Torsions Values					
Torsion No.	Value	S.D.	Torsion No.	Value	S.D.
1	58.43	23.82	22	9.40	179.51
2	19.87	172.14	23	1.55	174.29
3	-48.13	59.23	24	7.69	175.71
4	78.77	104.64	25	39.82	151.63
5	66.15	149.81	26	74.66	67.05
6	32.86	114.87	27	96.21	83.75
7	33.69	103.25	28	27.70	162.27
8	-39.90	147.14	29	-71.20	100.26
9	-55.64	74.84	30	67.65	118.09
10	-79.45	124.09	31	63.62	154.92
11	-35.68	163.10	32	30.67	162.17
12	4.81	175.67	33	60.36	149.04
13	-13.09	163.64	34	14.55	176.98
14	19.47	177.44	35	46.77	163.22
15	85.66	155.56	36	69.65	155.27
16	74.67	151.73	37	-24.29	175.43
17	15.20	175.89	38	10.09	176.02
18	3.89	173.43	39	12.04	179.95
19	-23.27	175.56	40	-20.31	174.61
20	-17.29	170.66	41	-131.11	44.09
21	-1.66	176.97			

A high percentage of the torsions possess S.D. values in the order of 100-180 and thus exhibit highly dynamic behaviour. Certain torsion angles clearly are an exception to this, having significantly lower S.D. values, and these are associated with either the choline or phosphate group.

Thus, this data again indicates that the acyl chains are flexible whilst the head groups are pseudo crystalline in structure.

5.2.1.7. Interface and Head Group Hydration

The Density Profile plots in Fig. 5.8 illustrated the level of water hydration and penetration into the bilayer interior and also shows a 10 Å of bulk water in the inter-bilayer region.

Fig. 5.10a contains the oxygen Radial Distribution Function (RDF) profile for the glycerol moiety, the Phosphatidyl moiety and for the water molecules, and Fig. 5.10b contains the RDF profile for the Choline entity.

The hydration differential that the DPF plots suggest is also reproduced in the glycerol (Gl) and phosphatidyl (PC) RDF plots for oxygen and hydrogen atoms. Firstly, the peak of the PC water 1st hydrogen shell is 9 times higher than the Gl peak, thus indicating that there is higher solvation around the phosphatidyl than the glycerol oxygens. Secondly the peak of the phosphatidyl groups 1st hydrogen shell is closer to the centre of the RDF than that of the glycerol 1st hydrogen shell. Thirdly, the phosphatidyl group induces clearly definable secondary and tertiary shells in the aqueous phase whereas the glycerol oxygens only induce a slight perturbation equivalent to the position of a secondary shell.

The first effect results from the higher water content in the upper bilayer region, around the PC moiety and the reduction in water density to zero in the hydrophobic interior. The second and third effects result from the reduced water penetration at the glycerol entity and the PC group having a long range effect on the surrounding water molecules due to its dipole.

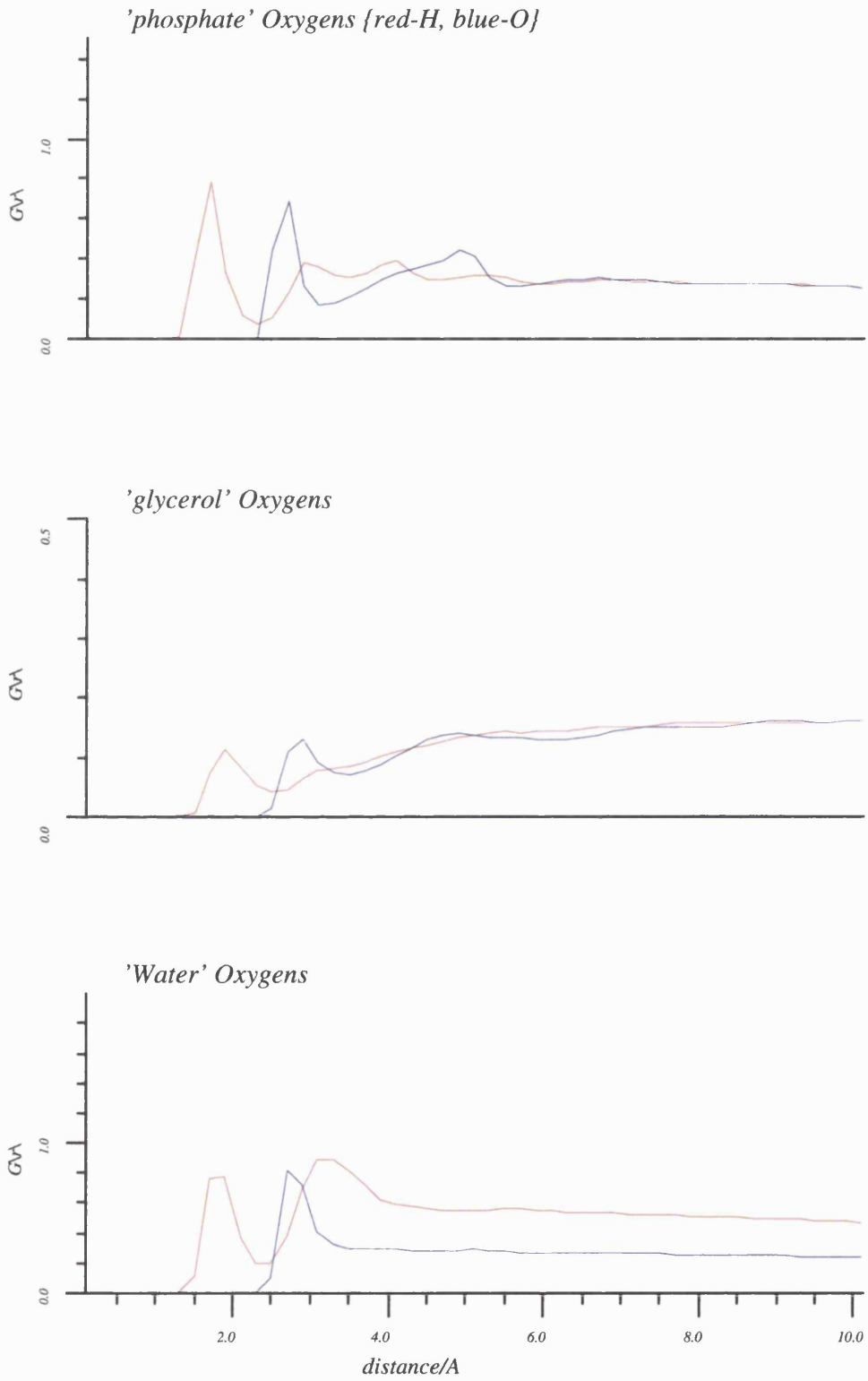


Fig. 5.10a Hex - Water Radial Distributions

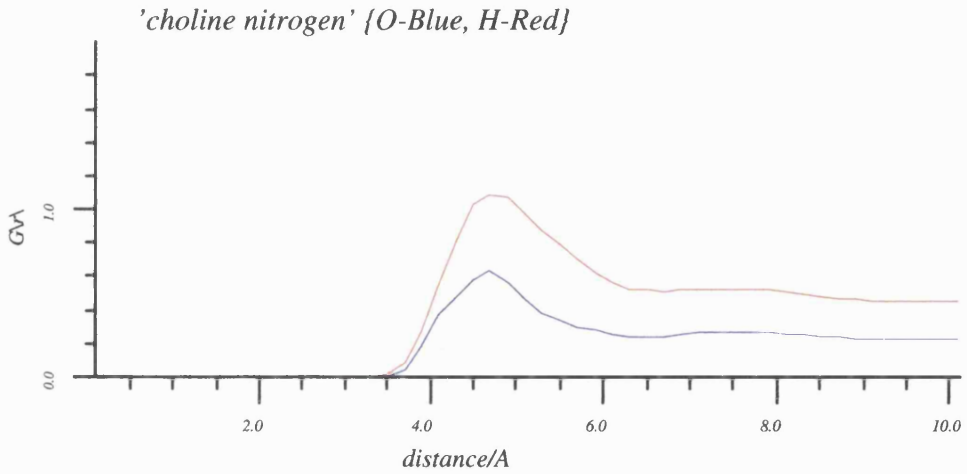


Fig. 5.10b Hex - Water Radial Distributions

The first PC water oxygen (O) shell can be seen at 2.7 Å with its associated hydrogen (H) peaks at 1.9 Å and 3 Å. The hydrogen peak at 1.9 Å equates to the aligned hydrogen, as the distance between the O and H is approximately equal to an O-H bond distance, and the peak at 3 Å equates to the non-aligned secondary hydrogen shell. The second water shell has its oxygen peak at 5 Å, with its aligned hydrogen peak at 4 Å and its intershell hydrogen peak at 5.4 Å.

The glycerol oxygen atoms have a first oxygen peak at 2.95 Å, with an aligned hydrogen peak at 1.95 Å, but no secondary hydrogen peak can be observed. Only a slight secondary oxygen peak can be observed at 4.75 Å.

Fig. 5.10b shows the RDF of water molecules around the nitrogen of the choline group. These values were calculated in order to investigate the water structure around the head group moiety and to look at whether the structure was determined by its bulky nature or its charged nitrogen.

Fig. 5.10b shows the hydrogen RDF to have a peak at 4.75 Å and a slight peak at

≈7.5 Å. Likewise the oxygen plot shows a water oxygen peak at 4.75 Å and a slightly discernible peak at ≈7.5 Å. Thus, the hydrogen and oxygen peaks are at the same displacement from the choline nitrogen, which indicates that the waters take up a clathrate style structure and thus, the bulky nature of the choline predominates over its charge. In the clathrate structure the waters satisfy their hydrogen bonding requirement through interaction with other water molecules and thus, they 'turn away' from the choline.[5] The only slight difference in the hydrogen and oxygen profile is a broadening of the H-peak on its back edge, which is probable a result of a repulsion between the positive charge on the choline nitrogen and the partial charge of the hydrogen of the waters.

5.2.2. Discussion on Hex System Bilayer Results

Fig. 5.11 shows the unit cell coordinates of the Hex system at 0 ps, 100 ps and 200 ps without the inter-bilayer water layer.

As mentioned earlier the build and refinement procedure attempted to minimise the level of crystallinity in the acyl chains and to reduce the potential co-operative motion early in the simulation. A visual inspection of parts 2 and 3 of the Fig. 5.11 indicates that the bilayer has a rough surface with randomly orientated lipids and fluid acyl chains. The degree of randomness in the bilayer can be seen to increase as the simulation progresses.

The thermodynamic data (Fig. 5.7) and the average temperatures (Table 5.3) indicate that the system is equilibrated. The UCV plots in Fig. 5.4 show that the unit cell is relatively stable under NPT conditions and therefore there is justification[12] for the use of NVT conditions for the remainder of the simulation in order to maintain the required SA and to reduce the structural fluctuations that is produced by the pressure bath methodology.

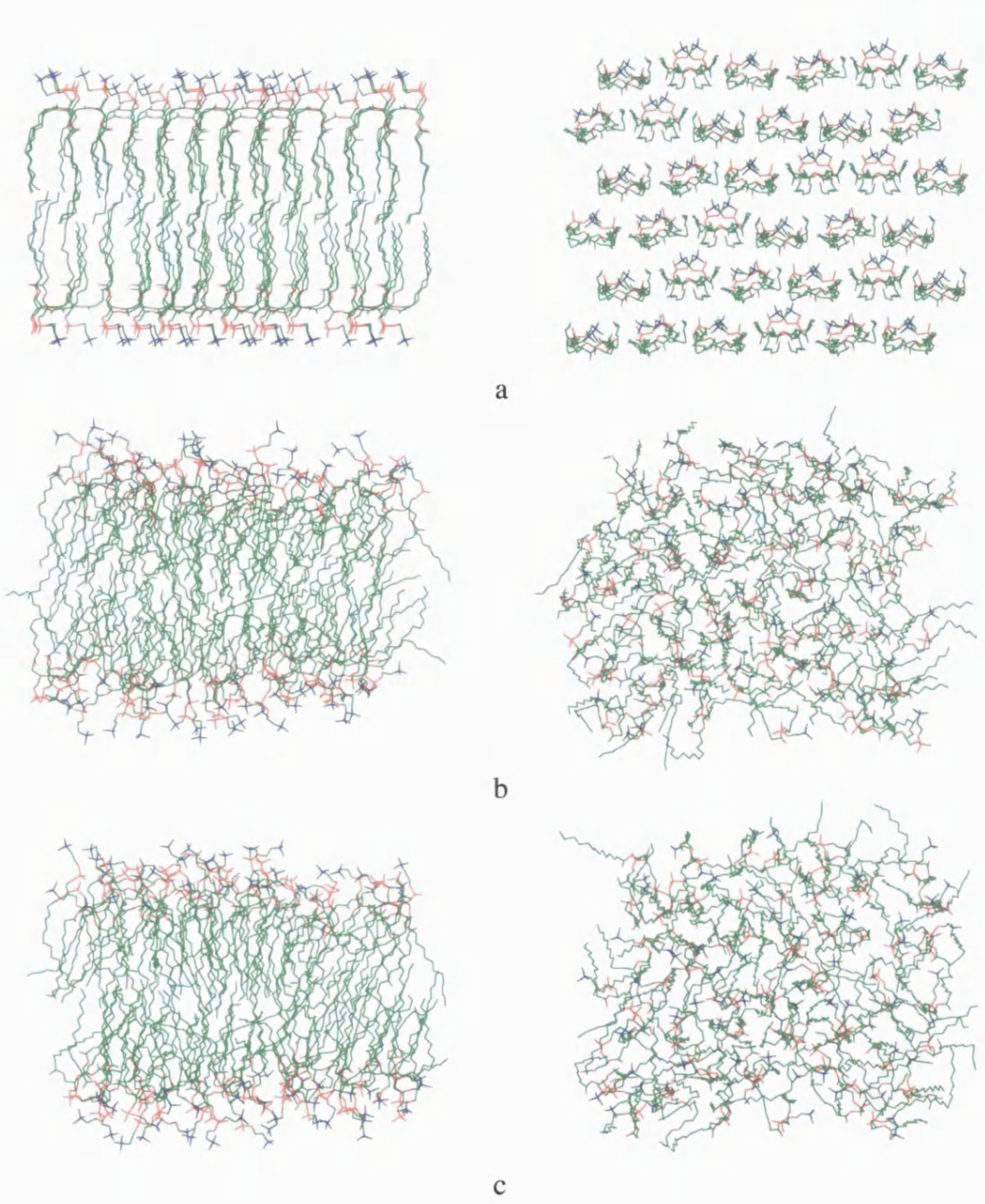


Fig. 5.11 Hex - Unit Cell Coordinates at 0 (a), 100 (b) and 200 ps (c)

The DPF results (Fig. 5.8) and the RDF plots (Fig. 5.10) indicate that the structure of the bilayer is typical of a good model DMPC bilayer, reproducing the Marrink and Berendsen model of the bilayer heterogeneity, the rough surface, the differential hydration around the phosphatidyl and glycerol oxygens, the long range head group

perturbation of the water layer, the clathrate water structure around the choline group and the 10 Å bulk water phase.

The lipid structural lengths (Table 5.4) and angles (Table 5.5) show some interesting results, whilst indicating that the head group is a pseudo-crystalline entity and the two acyl chains are fluid, as observed experimentally, the data also shows a differential between the Sn1 and Sn2 acyl chain dynamics. This can be rationalised by the different geometry of the 2 chains as a result of the bend in the Sn2 chain near to its junction with the glycerol and the flexibility of the glycerol-acyl chain torsions. The SOP plots (Fig. 5.9) show good agreement with experimental data in both the profile and in the values of the order parameters. The torsion data (Table 5.7) shows again that the chains are fluid as their S.D. values are relatively high but the t/g conformation populations (Table 5.6) indicate that whilst the torsions are pretty flexible they do not undergo a high number of conformation changes.

Therefore, the hexagonally packed, fully hydrated DMPC bilayer possess a mobile, fluid acyl chain region with a more crystalline head group region, but with the acyl chain fluidity arising from a combination of t/g isomerisations and scissoring of the chains. This relatively reduced level of isomerisation could be a result of the warming process utilised to attempt to minimise any high energy interactions induced by the UCV scaling of the system or an inaccuracy in the potentials describing the acyl chain methylene groups. Through the refinement process some of the torsional energy held in the pre-equilibrated acyl chains could have been lost.

Thus, these results indicate that the DMPC hexagonal bilayer reproduces the majority of the bilayer stability indicators and could confidently be used as a model to study other bilayer properties.

5.2.3. Cubic Packed Bilayer -Sq

5.2.3.1. Average Temperatures

The atomic temperature plots for the atoms comprising the different sub-sections of the bilayer system are illustrated in Fig. 5.12.

As with the hexagonal packed bilayer simulation the use of a multiple temperature bath resulted in the atomic temperatures being near the required average of 311 K. The average values are contained in the Table 5.8.

Table 5.8

Sq, Average Temperatures (K)		
Group	Temperature	S.D.
Sys	312.3	1.6
Chol	311.8	5.1
Gl	313.8	9.3
Sn1	311.1	4.0
Sn2	309.8	4.2
Wtr	314.3	2.9
Acyl	310.5	2.6

Again, as with the Hex system the average values are above or below the temperature of the multi temperature bath. This follows from the explanation discussed earlier for the Hex simulation.

The other main feature of the temperature plots is the warm up phase during the refinement of the structure which is seen in a series of steps as the temperature is raised at the end of each 1 ps period.

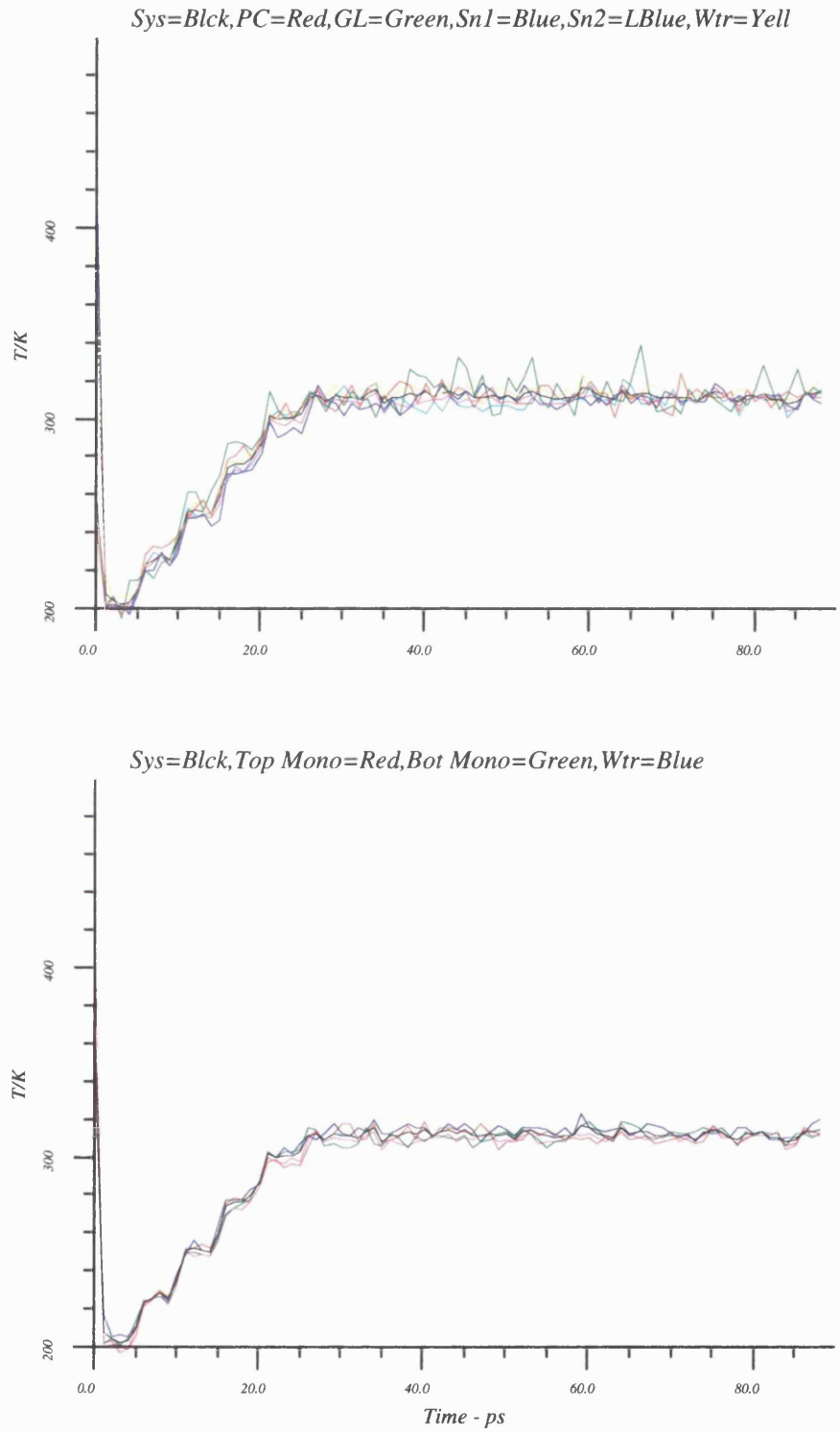


Fig. 5.12 Sq - Average Temperatures

5.2.3.2. Thermodynamic Data

The thermodynamic data is shown in Fig. 5.5.

The Unit Cell Vectors (UCV), as with the Hex system, stabilise following the UCV scaling and the NPT conditions are applied. The thermodynamic energies are also stable at the end of the scaling procedure, and after the transition to NVT conditions. Again the UCV are stable following the application of the NVT conditions at 35 ps.

The only change in the thermodynamic data resulting from the application of NVT conditions is a relative lowering of the pressure to a more negative value, which is not observed in the Hex system. This indicates that density of the system is a little below that required for the VFF and thus, the UCV values were a little high. But these values for the UCV were relatively stable in the NPT system so there could be another reason for the drop in pressure after this point in the simulation other than just a density change induced by the NPT conditions.

5.2.3.3. Density Profile Function Data

The density profiles (DPF) for the sub sections of the bilayer are illustrated in Fig. 5.13 for the Sq system.

Comparing the form of the Sq simulation DPF profile to that of the Hex system it can be clearly seen that the profile is more uneven, e.g. the water profile in the Hex system simulation has an even profile but in the Sq system it has an undulating profile. This is due to the short length of this run, which does not allow the smoothing of the built structure and the relaxation of the different groups to a more accurate and diffuse profile. This can also be seen in the Phosphatidylcholine (PC) peak compared to the wider, broad based Hex PC profile. The partitioning of the densities is in general agreement with the Hex system

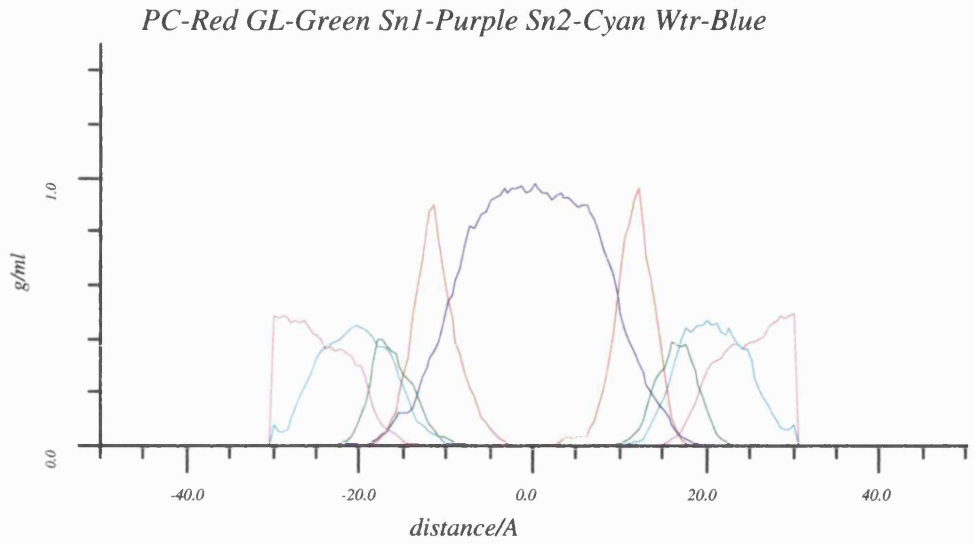


Fig. 5.13 Sq - Density Profile

5.2.3.4. Lipid Structural Internals

Table 5.9 shows the average values and standard deviation (S.D) for the selected structure angles that describe the internal motion of the lipid molecules in the Sq simulation.

Table 5.9

Sq, Lipid Structural Angles		
Number	Value	S.D.
θ_1	106.92	11.91
θ_2	36.17	19.51
θ_3	95.35	23.40
θ_4	121.74	19.21
θ_5	101.01	18.74
θ_6	134.73	22.32

The S.D. values indicate that the lipid angles have reasonable mobility, but there is not a clear pattern as seen in the S.D. of the Hex. The short length of this simulation may mean that the disparity in the motion of the chains is yet to develop or the lack of

this mode of motion could be a result of the geometry of the lipids. It is also clear from a comparison of the two sets of angle data that there is much higher flexibility in the Hex system than in the Sq system.

All the angle values are similar, except θ_1, which describes the motion of the two acyl chains relative to the glycerol carbon and is 10° below that of the Hex system. This could indicate that the two lipid chains remain in a more aligned structure.

Thus, the data indicates that the acyl chains chains are less mobile than those of the Hex system and has a very much reduced motion from the 'scissoring' action. θ_1 indicates that the head group is along the bilayer surface but with a higher component into the water layer than the Hex system.

Table 5.10 shows the average vales and S.D for the selected structure lengths that describe the internal motion of the lipid molecules in the Sq simulation.

Table 5.10

Sq, Lipid Structural Lengths		
Number	Value S.D.	
d_1	4.67	0.22
d_2	6.77	0.64
d_3	17.64	1.41
d_4	15.05	1.39
d_5	13.96	0.91
d_6	13.84	0.80
d_7	9.81	3.51
d_8	38.09	2.46

The structural lengths have a similar form to that of the Hex system data, but a difference again is seen in the relative mobilities of the Sn1 and Sn2 chains. The S.D. values for the Sn1 and Sn2 acyl chains are more similar than in the Hex simulation indicating a more common dynamic profile. The Sq acyl chains show lower mobility and a 3 Å lower inter-termini spacing than the Hex simulation.

The data once more indicates that there is a difference in the motion of the two

acyl chains between the systems, resulting from the rigid body scissoring action of the two chains and is reduced in the Sq simulation.

d_8, the P-P distance, during this simulation shows little or no change has occurred to the values. This indicates that the short MD simulation has not resulted in a value close to the experiment.

5.2.3.5. Torsional Data

It can be clearly seen from Table 5.11, containing the lipid torsional standard deviation values and Table 5.12, containing the trans gauche (t/g) acyl chain torsional percentages, that the Sq system simulation has a very similar profile to that of the Hex system simulation.

Table 5.11

Sq, Lipid Torsion Values					
Torsion No.	Value	S.D.	Torsion No.	Value	S.D.
1	61.27	15.64	22	37.71	174.54
2	177.11	25.15	23	107.11	139.01
3	-46.85	57.90	24	48.07	165.04
4	62.10	60.25	25	31.99	147.94
5	122.80	44.61	26	83.24	26.85
6	54.96	66.86	27	119.54	34.33
7	72.22	34.90	28	89.43	141.84
8	-89.61	91.98	29	-59.45	110.33
9	-67.56	48.54	30	96.92	41.74
10	-132.41	66.87	31	112.30	88.34
11	-112.72	128.02	32	-17.60	183.52
12	29.21	173.70	33	90.93	137.78
13	34.81	178.58	34	-6.96	179.19
14	-11.87	178.94	35	-24.04	154.07
15	86.67	143.59	36	-38.78	172.40
16	123.46	104.33	37	-16.92	158.38
17	10.04	180.14	38	22.43	180.41
18	-40.82	169.67	39	14.52	176.31
19	-12.86	168.44	40	-150.30	92.56
20	-2.01	179.68	41	-135.68	40.96
21	-15.92	177.87			

Table 5.12

Sq, Lipid Acly Chain t/g Percentage							
Sn1				Sn2			
Torsion No.	t	g+	g-	Torsion No.	t	g+	g-
14	85.26	0.43	11.10	29	24.89	1.81	66.92
15	82.14	2.41	12.58	30	58.05	0.96	37.59
16	81.34	3.67	11.95	31	78.58	3.03	14.71
17	88.20	2.81	6.55	32	80.77	3.73	11.95
18	81.41	5.90	9.27	33	85.05	3.01	8.83
19	78.38	7.10	11.20	34	74.50	4.68	17.39
20	81.18	3.72	11.70	35	88.56	1.84	6.94
21	93.06	1.51	3.32	36	76.13	10.84	9.79
22	92.00	0.99	4.93	37	90.90	4.43	2.63
23	83.48	2.70	11.15	38	86.60	4.91	6.20
24	89.55	3.50	4.73	39	88.14	4.01	5.66
Ave.	85.10	3.16	8.95	Ave.	75.65	3.93	17.15

The number of gauche torsions per chain is 1.10 for the Sn1 and 1.91 for the Sn2 chain is slightly higher than that for the Hex simulation. The cause of this result will be discussed later in the Chapter.

5.2.3.6. Interface and Head Group Hydration

The water radial distribution (RDF) plots for the phosphate, glycerol and water oxygens are illustrated in Fig. 5.14a and for the choline oxygens is illustrated in Fig. 5.14b.

The form of all three plots is very similar to that of the Hex system plots, with the water shells being at the same displacements from the origin. The main differences are firstly, that the PC peaks are a little less defined and broader, as the shell structure is still developing. Secondly, the peaks in the glycerol oxygen RDF are lower that those in the Hex system simulation, indicating reduction in the degree of water penetration. All these are probably due to the system still equilibrating and to observed differences in bilayer structure.

The water RDF for the choline nitrogen atoms are illustrated in Fig. 5.14b. The plot again shows agreement, in the peak positioning and general form, to that of the Hex system but with evidence of features not yet fully developed, e.g. a slightly higher peak value (probably due to an unequilibrated water phase) which is consistent with the DPF results. Secondly, the slight repulsion of the water hydrogens by the large $-N(CH_3)_3$ dipole, causes a displacement of the peak.

Thus, the RDF data shows that a good water structure is beginning to develop in this simulation and it replicates most of the required results including the clathrate structure around the head group.

5.2.4. Discussion on Cubic System Bilayer Results

All the evidence above indicates that the bilayers derived from the cubic packed geometry possess some of the features required in a good model bilayer. Obviously, these features are only beginning to develop at this relatively short time frame.

Fig. 5.15 shows the cubic system coordinates at 0 ps and at 90 ps. On first inspection the bilayer looks fluid, with a rough surface, indicated by the uneven penetration out of the bilayer of the head groups. A closer inspection of the coordinate highlights a tilt of the lipids is beginning to develop, which is a feature usually associated with the $L_{\beta'}$ bilayer phase and not with the physical L_{α} phase. The tilt of the bilayer structure occurs because the acyl chains slant from the top right to the bottom left of the unit cell. The system and group temperatures are near the required value for the L_{α} phase so the tilt is not a temperature related problem.

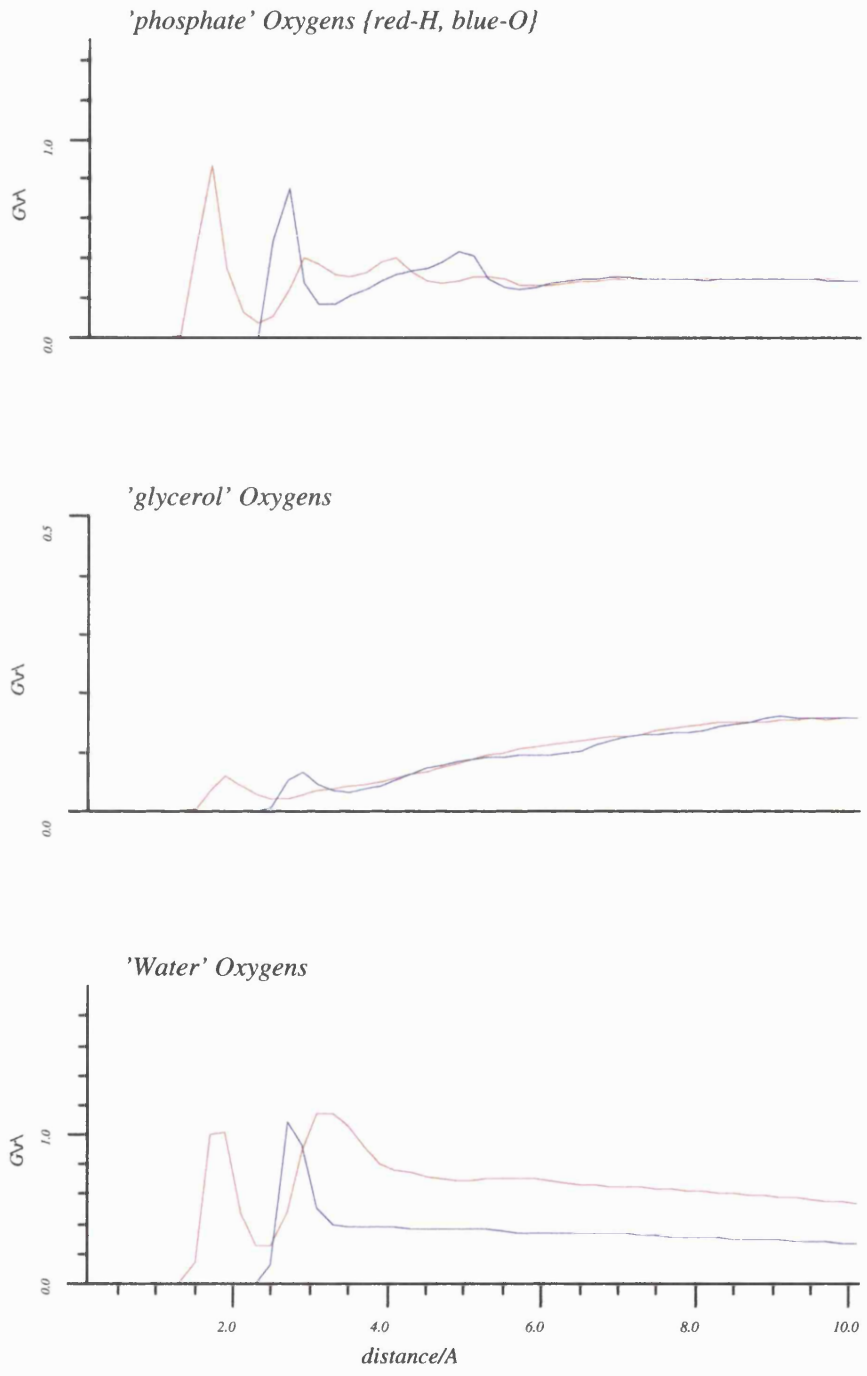


Fig. 5.14a Sq - Water Radial Distribution

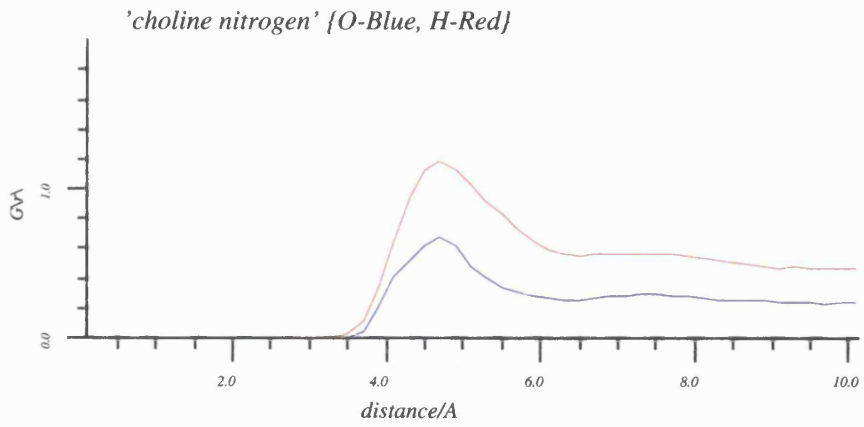


Fig. 5.14b Sq - Water Radial Distribution

As mentioned during the building process, the cubic packed geometry has a more aligned structure, with the acyl chains in a position to act more cooperatively than the psuedo randomised Hex structure (Fig. 5.1). During the early stages of the simulation the chains must have acted cooperatively and tilted to take up the space between themselves and not through the t/g isomerisation that is thought to be the mechanism in physiological bilayers. This tilting of the lipids also caused a reduction in the glycerol motion, as indicated by the structural data, and appears to be the reason that the scaling of the unit cell for the Sq system had no significant effect on its P-P distance and SA. The head group angle and the glycerol hydration indicate that the more dense Sq dynamic structure increases the inter-head group interactions, causing them to be more directed into the water layer. Also, the higher peak in the Sq RDF for the water-water oxygens reflects that the higher P-P distance in this system is causing the head group to interact more with the water layer.

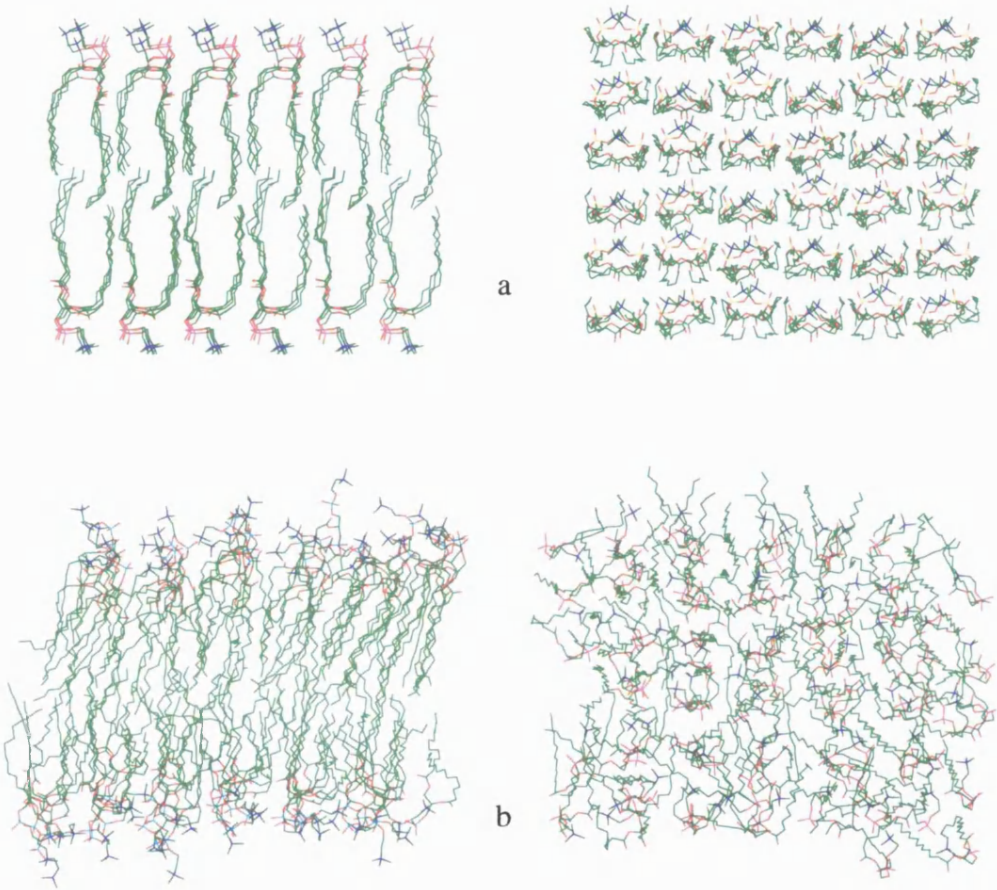


Fig. 5.15 Sq - Unit Cell Coordinates at 0 (a) and 90 ps (b)

5.3. Conclusion

In both the simulations an extensive analysis of the physical structure and thermodynamic properties of the two systems have been undertaken to try to answer two questions:

1. Can we build a bilayer model that replicates an experimental bilayer?
2. If so, which packing arrangement gave the best model bilayer?

Both systems reproduced general bilayer properties well but there were two main areas that distinguished the two:

1. The physical dimensions of the Sq system indicated a too high P-P distance.
2. The HC chains in the Sq system acted cooperatively to produce a tilted phase similar in look to the $L_{\beta'}$ phase, whereas the Hex system remained essentially fluid L_{α} phase.

Whilst, the Hex bilayer reproduced more of the required bilayer properties, including a fluid interior, it also had a lower than experimental gauche population per chain.

Table 5.13

Rate of Change of Acyl Chain t/g Population				
	Sn1		Sn2	
Time window (ps)	g%	-Δg	g%	-Δg
0-5	23.85		32.33	
5-10	23.71	0.14	31.78	0.55
10-15	22.10	1.61	29.88	1.90
15-20	20.73	1.37	28.85	1.03
20-25	19.13	1.60	27.82	1.03
25-30	18.55	0.58	26.18	1.54
30-35	17.46	1.09	25.36	0.82
35-40	16.72	0.74	24.90	0.46
40-45	16.54	0.18	24.55	0.35
45-50	15.58	0.96	23.86	0.69

Table 5.13 above contains the rate of change of the gauche populations (Δg) for the two acyl chains averaged over 5 ps time windows through the refinement part of

the building process. This was to done to investigate the origin of the reduction in the gauche population and from which part of the simulation the majority of the chain dynamic motion was lost.

From Table 5.13 it can be concluded that in the first 50 ps of the simulation Sn2 gauche population is lost, which is indicated by the Δg values. This points to the refinement process having a negative effect on the t/g percentage. In this procedure the density of the system is being reduced quite considerably to approach that in physiological bilayers. Such an increase in density is known to reduce the probability of alkyl chain isomerisation.[6] Another problem that could be effecting the population of gauche isomers in the acyl chains is the relatively high gauche torsion barrier employed in the V.F.F.. This is therefore an an important factor in the lost t/g percentage because following the refinement procedure a further 4.50% and 4.74% t/g was lost in the main simulation section, 50-220 ps.

As was discussed earlier, great care should be taken not to influence the simulation path with a poor starting structure, thus pre-equilibrated lipids were used in order to try to improve the success of the simulation. Unfortunately, whilst care was taken loss of t/g population still occurs.

Under the V.F.F. force field we have clearly been able to reproduce a high percentage of the results of other MD and experimental studies. Thus, it was decided that the bilayer derived from the hexagonally packed starting structure best represented the physiological bilayer and was therefore used as the starting point for the other bilayer studies described in Chapters 6 and 7.

References

1. Biosym. Technologies Inc. San Diego, CA, USA
2. A. Seelig and J. Seelig, *Biochemistry*, **13**, 4839-4845 (1974).
3. H. Heller, M. Schaefer, and K. Schulten, *Journal of Physical Chemistry*, **97**, 8343-8360 (1993).
4. S. J. Marrink, M. Berkowitz, and H. J. C. Berendsen, *Langmuir*, **9**, 3122-3131 (1993).
5. K. V. Damodaran and K. M. Merz, *Langmuir*, **9**, 1179-1183 (1993).
6. A. P. Lemon. Ph.D. Thesis Bath, 1992
7. S. J. Marrink and H. J. C. Berendsen, *Journal of Physical Chemistry*, **98**, 4155-4168 (1994).
8. A. J. Robinson, W. G. Richards, P. J. Thomas, and M. M. Hann, *Biophysical Journal*, **67**, 2345-2354 (1994).
9. K. Tu, D. J. Tobias, J. K. Blasie, and M. L. Klein, *Biophysical Journal*, **70**, 595-608 (1996).
10. G. Cevc and D. Marsh, *Phospholipid Bilayers,* Wiley-Interscience, New York (1987).
11. J-P. Douliez, A. Leonard, and E. J. Dufourc, *Biophysical Journal*, **68**, (5), 1727-1739 (1995).
12. D.P. Tieleman and H.J.C. Berendsen, *Journal of Chemical Physics*, **105**, (11), 4871-80 (1996).

6. Single Helix Incorporated Bilayer Simulations

6.1. Building The Helices

A poly-alanine peptide was built as both a template for the production of the other peptides used in this study and as a reference, control pure alanine helical peptide. The program MOLEDT [1] was used to build the initial primary sequence which contained 30 L-alanine residues and was blocked at the N-terminus by an acetyl group and at the C-terminal by a methylamide group. The blocked version of the peptide was used so that the helix represented the situation in a protein helix where the transmembrane section is linked to another section of the protein and thus, does not have charged termini. Such charged termini will have an effect on the helix backbone stability. The 30 residue peptide was chosen as its length was close to the distance spanning the bilayer hydrophobic interior and interface as observed in the pure bilayer simulation, Hex. Also, protein TM helices, by their very nature, span the entire bilayer as they are connected to non-membrane bound sequence.

MOLEDT was then used to impart a right handed α helical structure on the above 30 residue alanine peptide and the Steepest Descents minimiser was used to refine the backbone hydrogen bonding network and optimise the structure in the VFF.

MOLEDT was used to mutate designated residues in the alanine helix to selected amino acid residues. The residues chosen represented the main groups of the natural amino acids described in the Chapter 1. The three residues chosen for mutation were at positions 4, 17 and 29. These particular residues were chosen to facilitate the investigation of the effects that the different domains in the bilayer have on the mutated helices - residues 4 and 29 will give an insight into the hydrophilic head group region and its associated solvating water molecules whereas residue 17 will provide an insight in the dynamic behaviour of the mutated amino acids in the hydrophobic bilayer interior.

Fig. 1.2 shows the amino acids (Threonine, Arginine, Valine, Phenylalanine, Apartic acid, Aspartamine) used in the simulations described in this chapter. Figure 6.1 shows two views of the pure alanine helix, along and down the helix axis.

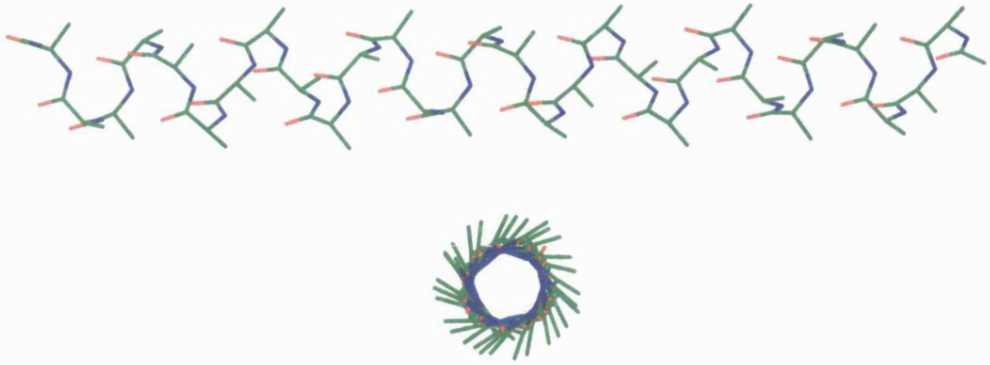


Fig. 6.1 - 30-mer Blocked Alanine Peptide α_R Helix

6.2. Building The Bilayer Receptor Cavity

The DMPC bilayer structure derived from the hexagonally packed geometry simulation in Chapter 5 (Hex) was used to construct a bilayer cavity that would accept the described single helices.

The first attempt to incorporate the alanine helix involved removing a centrally located pair of lipids. Due to the fluidity of the acyl chains of the molecules adjacent to the removed pair, the available space was restricted. Thus when the peptide was placed in this type of cavity catastrophically high non-bond energies were produced, which destroyed the integrity of the helix structure.

In order to produce a better cavity in the centre of the bilayer it was decided to use a spiral of fake atoms with an initially small Van der Waals (VdW) radius profile. The VdW radii of the fake atoms were gradually increased over 20 ps, in a normal VFF simulation, with conditions applied as per the Hex simulation. The initial value of the VdW parameters was close to those of hydrogen and over the 20 ps period the parameters were changed every pico-second until equal to that of the alanine helix, a

cavity diameter of ≈8-9 Å. This gradual increase in the profile of the spiral slowly repelled the acyl chains out of the required cavity.

Figure 6.2 shows the cavity after the removal of the lipid pair with the spiral of fake atoms inside it. The second graphic shows the cavity at the end of the process, showing clearly the clean nature of the cylindrical hole.

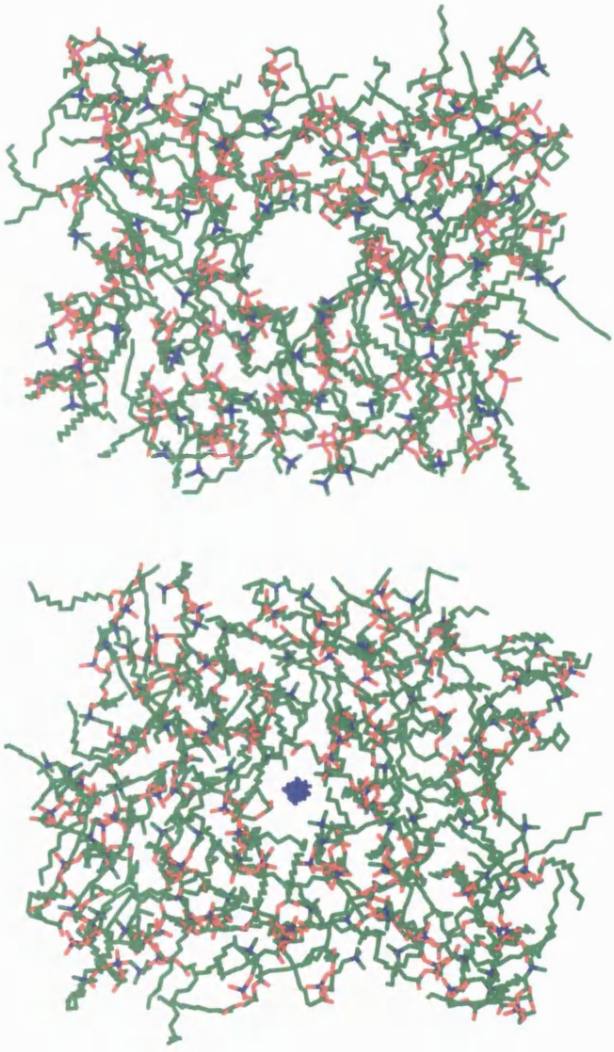


Fig. 6.2 - The Bilyer Cavity Before(Bottom-including spiral atoms) and After(Top) Refinement.

During this process it was important that the surface area occupied by the spiral was monitored in order that the value did not exceed that of the lipid pair removed, $\approx$65 Å^2. After the first 14 ps of the simulation, the radius of the spiral was such that the cross-sectional area equaled that of a lipid head group surface area (SA). At this point it was discovered that even though the Van der Waal volume of the spiral was below that of an ideal alanine helix the acyl chains had moved sufficiently to allow the helix to be inserted without high energy clashes occurring. Thus, the simulation was stopped at this point.

The non-bond energy was also monitored during this process to follow the effect that the repulsion of the acyl chains had on the system. The non-bond energy rose through this process, but the rise was well within that expected for such a process. Once the cavity was found to be at the required dimensions the simulation was carried on for another 30 ps to allow the system to settle down before the helical peptides were introduced.

During the above process the water layer was retained so that the bilayer remained as near as possible to its structure from the pure hydrated bilayer simulation, Hex.

After the cavity production procedure, the α_R helical peptides were placed in bilayer with their long axis along the bilayer normal (y-axis) and with residue 17 in the centre of the bilayer acyl chain region. This places the residues 4 and 29 in the interface region of the bilayer as required. Each mutant peptide was placed in the pre-equilibrated bilayer in turn, and again the water layer was retained so as not to add to the equilibration phase of this system.

6.3. Simulation Details

6.3.1. Minimisation

The seven bilayer systems:

Peptide	System Name
$ACE(ALA)_{30}N\text{-}M$	**Hex_Ala1**
$ACE(ALA)_{27}(ASP\text{–})_3N\text{-}M$	**Hex_Asp-**
$ACE(ALA)_{27}(ASP)_3N\text{-}M$	**Hex_Asp**
$ACE(ALA)_{27}(ARG)_3N\text{-}M$	**Hex_Arg**
$ACE(ALA)_{27}(PHE)_3N\text{-}M$	**Hex_Phe**
$ACE(ALA)_{27}(THR)_3N\text{-}M$	**Hex_Thr**
$ACE(ALA)_{27}(VAL)_3N\text{-}M$	**Hex_Val**

were then minimised to remove any clashes introduced during the above process. During the minimisation the coordinates of the helices were fixed in space to their original values, in order that the hydrated lipids did not adversely affect the helical structure.

The systems each had a small number of water clashes but the steepest descents minimiser removed these easily, thus essentially maintaining the original and pre-equilibrated structure. The steepest decents algorithm was used for as short a period as required to remove the higher energy interactions introduced in the build process. Table 6.1 shows the number of steps used in each of systems and the molecules to which the minimiser was applied. The only difficult minimisation involved the bulky and long chained arginine residue. The strain induced during the build and the water-peptide clashes were quickly removed but the strain caused by the side chains of the peptide proved difficult to remove. Thus, the minimiser was selectively applied to the clashing lipids around the cavity for a further 60 steps after which the whole system was minimised initially for 100 steps. This removed most of the strain in the cavity

region but it was necessary for the whole system to be minimised for a further 100 steps in order to remove residual strain.

Table 6.1

Minimisation details		
Sys.	nsteep	Molecules included
Hex_Ala1	20	all
Hex_Arg	100	all
	60	Cavity lining
	100	all
Hex_Asp	20	all
Hex_Asp-	20	all
Hex_Thr	50	all
Hex_Phe	50	all
Hex_Val	50	all

6.3.2. Molecular Dynamics Details

The 7 minimised systems were then put into the VFF molecular dynamics program and run for over 100 ps using a 1 fs time step and recording of the velocities and coordinates every 20 fs. A truncated non-bond interaction cut-off was used for the Lennard-Jones and coulombic interactions at 16 Å, with an 18 Å cut-off being applied to the neighbour list generation. Periodic boundary conditions were used, with the unit cell dimensions maintained from the building process above.

The initial system temperature for the random Boltzmann distribution of velocities was set at 200 K and raised gradually to 311 over the first 5 ps. The coordinates of the helices were restrained, by the use of a forcing constant, to their original values during the same initial 5 ps. Unlike the fixed method used in the minimisation section the coordinates of the helix are allowed to move but a RMS (root mean square deviation of the coordinate from its reference coordinate) retaining force adds an energetic penalty to the system, the higher the displacement the higher the penalty. This was

done to maintain the structure of the helices during the period of the simulation where the large input of kinetic energy could cause abnormal fluctuations in the dynamic behaviour of the system - any potential disruption to the helix at this early stage could bias the simulation path away from that directed by the force field and therefore result in the development of artefacts.

During the initial 5 ps period the velocities were reassigned every 1 ps in order to improve the radomisation and equilibration of the initial system.

6.4. Bilayer System Results

An extensive analysis of the bilayer structure has been carried for the seven single helix simulations. This analysis was identical to that for the pure bilayer simulations reported in Chapter 5. The structure of the peptides in these simulations was analysed and gives a good insight into the dynamics of such helices when incorporated into a bilayer environment.

The thermodynamic and temperature data are included below in order to quantify the equilibration of the systems. The important and interesting bilayer and helix structural results are included in sections 6.5 and 6.6, whilst the other remaining data is included in Supplementary Work, Section I.

Table 6.2

Hex_Ala1, Average Temperatures (K)		
	Temp	S.D.
System	312.2	1.6
PC	314.6	5.8
GL	313.0	9.0
Sn1	309.4	4.1
Sn2	309.9	4.0
Water	313.8	2.6
HC	309.6	2.7
Top Mon	311.4	3.4
Bot Mon	311.6	3.3

6.4.1. Average Temperature - Hex_Ala1

Fig 6.3 contains the group temperatures across the simulation for the Hex_Ala1 system. The different sub system groups were separately calculated from the sum of the individual component atomic velocities. Table 6.2 contains the simulation average values and the standard deviations of the averages (SD).

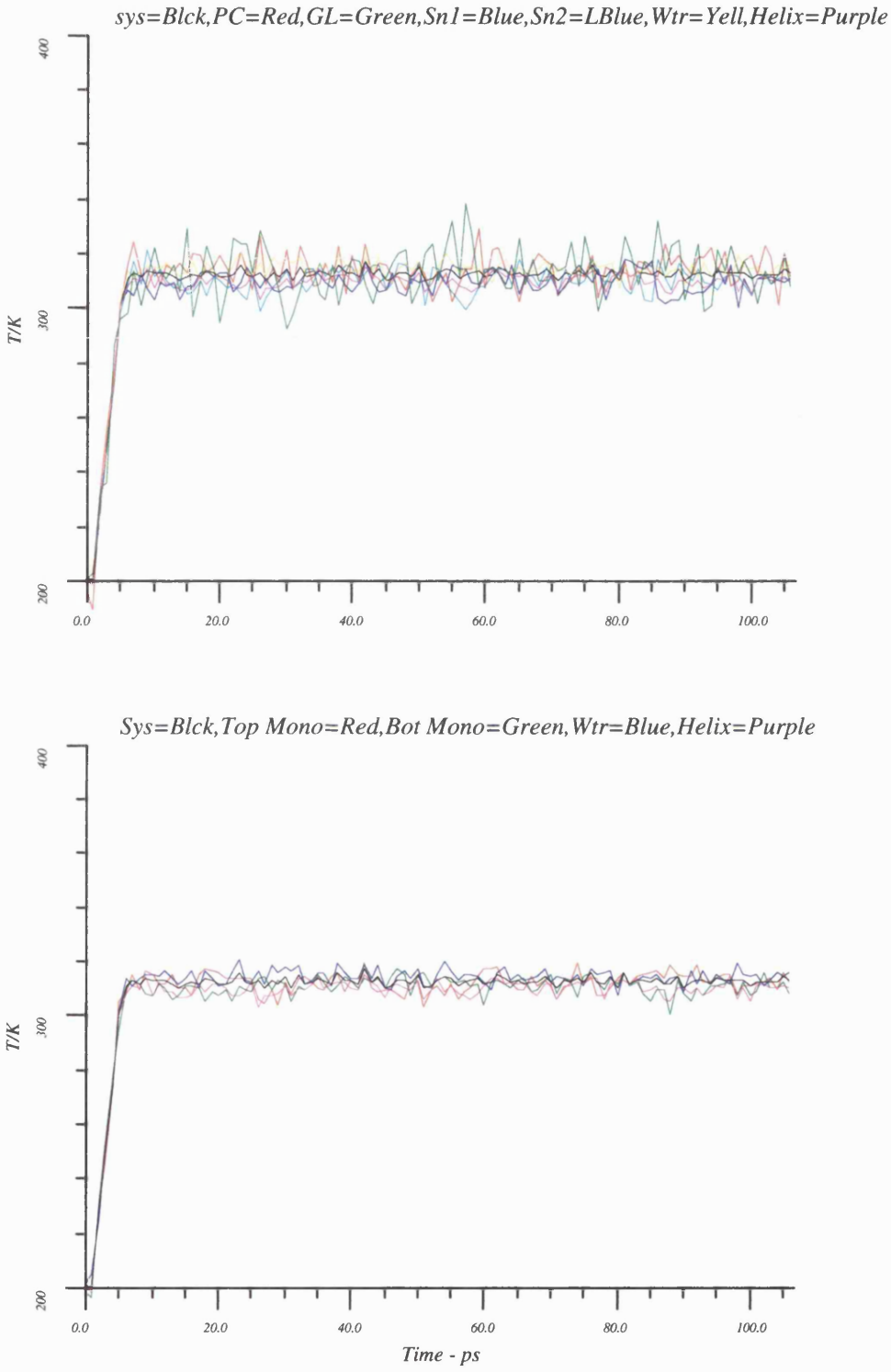


Fig. 6.3 - Hex_Ala1 Temperatures

6.4.2. Thermodynamic Data - Hex_Ala1

The general system conditions and the thermodynamic data for the Hex_Ala1 system are illustrated in Fig 6.4.

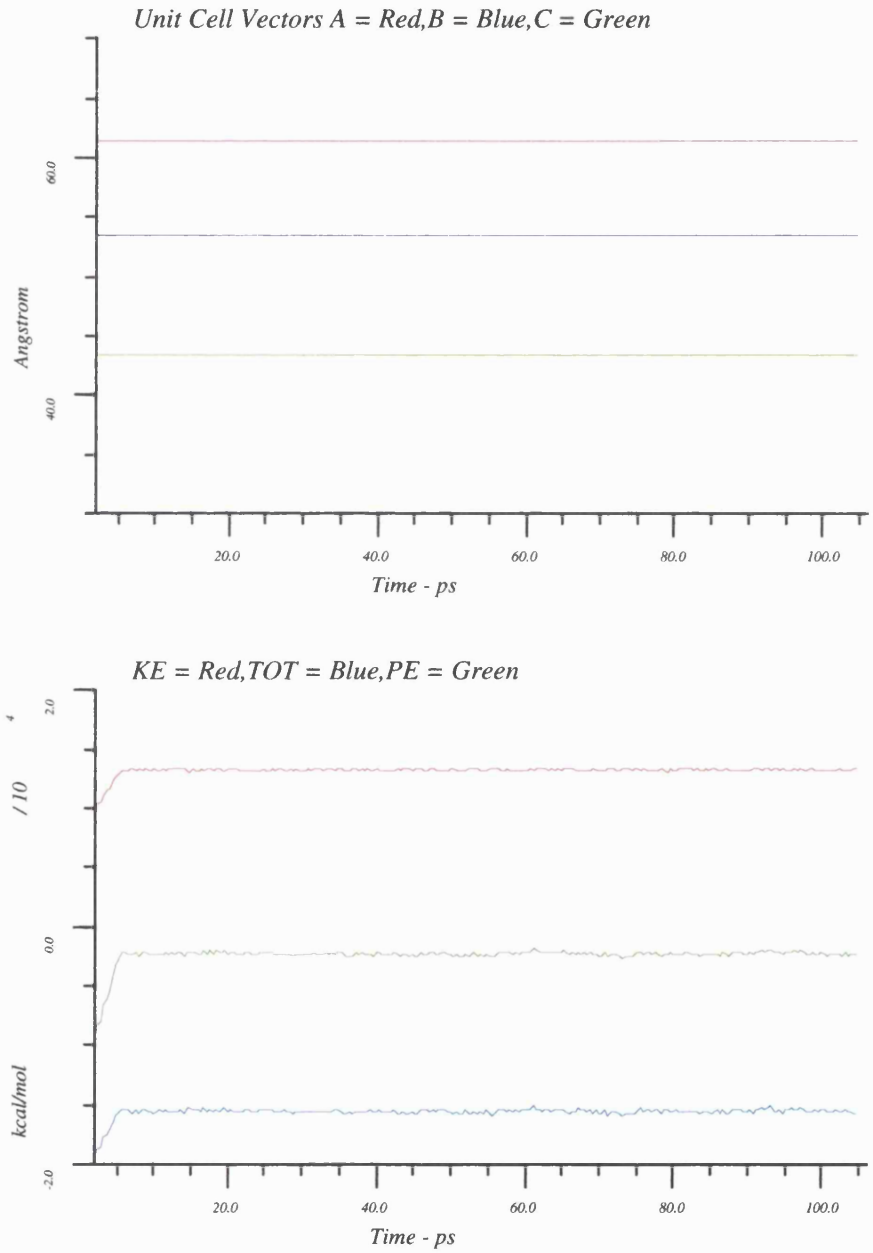


Fig. 6.4a - Hex_Ala1 Thermodynamic Data

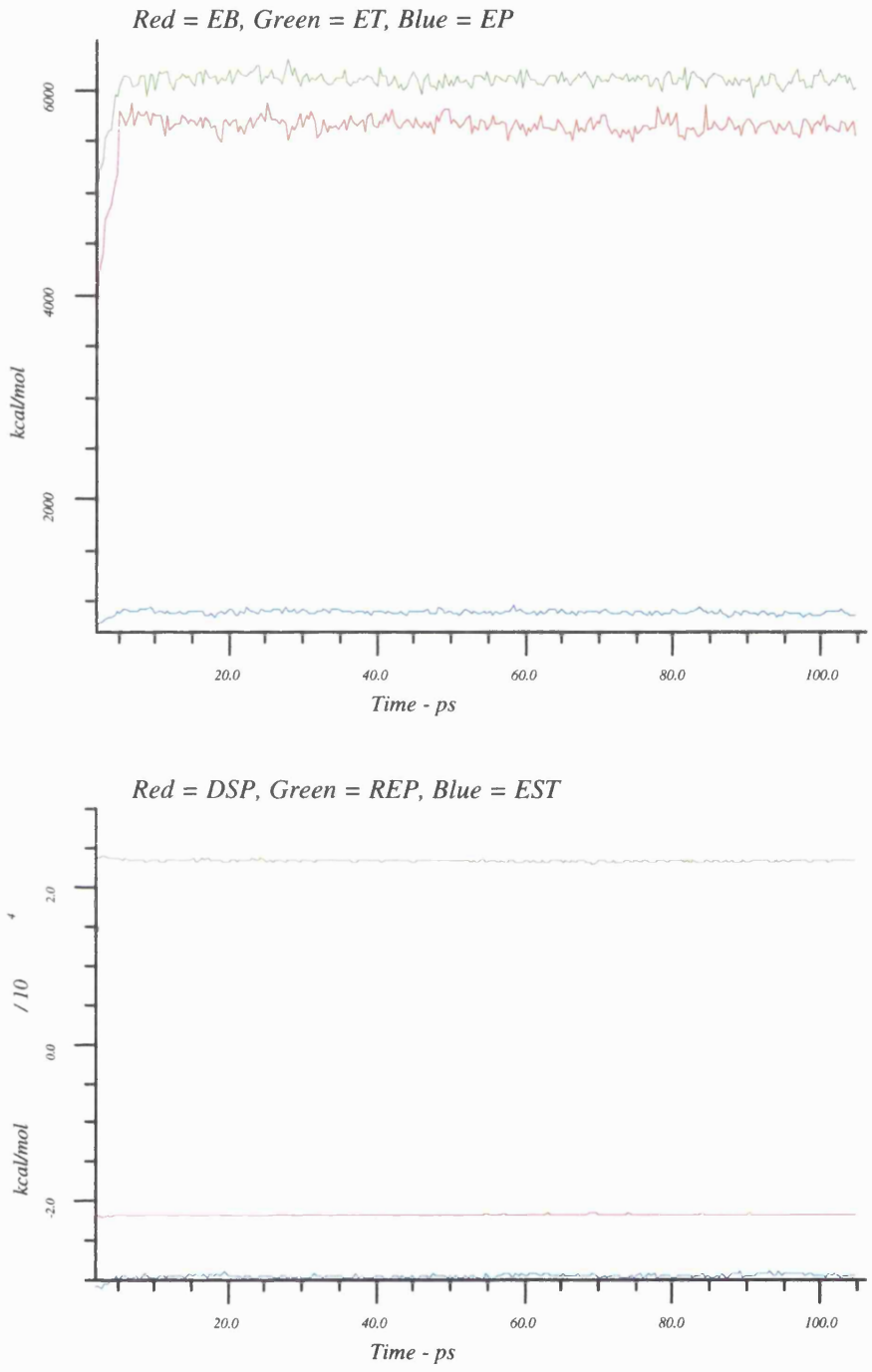


Fig. 6.4b - Hex_Ala1 Thermodynamic Data

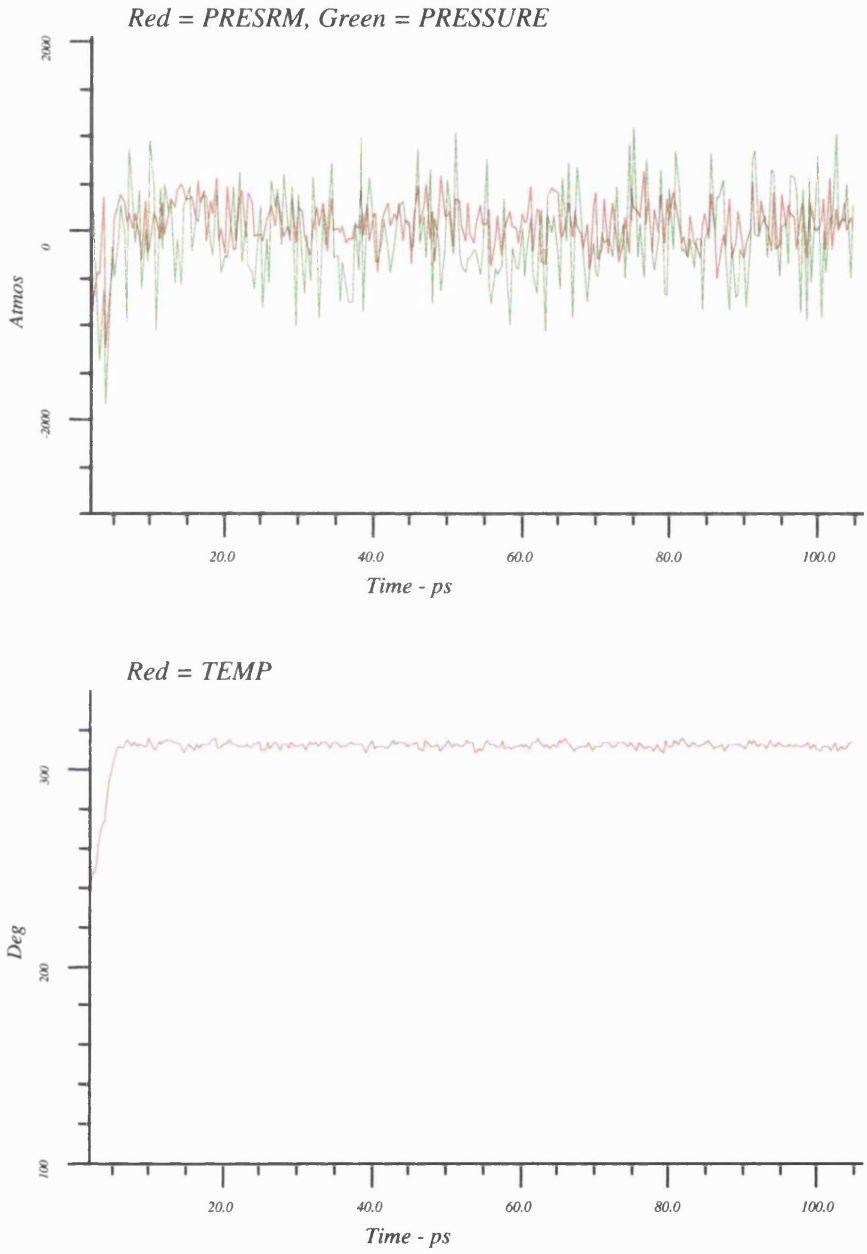


Fig. 6.4c - Hex_Ala1 Thermodynamic Data

As can be seen from the data, the average temperatures of the Hex_Ala1 simulation vary little from those of the pure bilayer simulation, Hex (Table 5.3). The cut-off induced temperature effect is clearly present again, as discussed earlier, with the more polar groups having an average slightly above the required value and the non-polar groupings having an average just below the required value.

The data for the Hex_Ala1 temperatures are typical of the whole set of single helix incorporated bilayer simulations (see Supplementary Work for detail).

From the energy plots in Fig 6.4a it can clearly be seen that the total, kinetic and potential energies are very stable throughout the simulation, after the initial short period of increasing thermal energy. This initial phase is where the system temperature is being increased gradually in order to try to maintain the integrity of the helix. Also, the Unit Cell Vectors (UCV) are constant as the NVT conditions are applied. Again the plots show the influence of the "warm-up" phase at the start of the simulation.

The torsion energy remains stable for the first 5 ps of the simulation and then shows a near instantaneous change to a new, stable value, when the RMS template forcing of the helix is released. The releasing of the helix restraints has no affect on the bond and theta energy. Thermal energy appears to take time to partition into the low energy dihedral angles.

The Van der Waals energy components (dispersive and repulsive) show good stability throughout the simulation whereas the electrostatic energy shows a change associated with the warm-up phase but is then stable to the end of the simulation. This obviously is a consequence of the increased motion of the helix and the subsequent effect on the backbone hydrogen bonds which directly affects the electrostatic energy, as does the increased motion of the interface lipid atoms and the water molecules.

The pressure also indicates a good degree of stability in the system, but as explained earlier they exhibit a degree of variability. This is also true for the system

temperature, which oscillates around a value that is stable after the initial warm-up phase.

Therefore, all the thermodynamic data for the Hex_Ala1 simulation indicates that the system is well equilibrated.

Comparing the Hex_Ala1 simulation data to the other single helix incorporated systems results indicates that they are all very similar (data not included). The similarity in the data thus suggests that the bilayers in these simulations are closely related thermodynamically and they are all stable. This is as would be expected for systems that include a pre-equilibrated bilayer and are closely related structurally, as here, where the only major differences are the three mutated residues.

6.5. Bilayer Structural Data

The incorporation of the helices into the bilayer appears to have caused significant changes in the bilayer structure. Firstly, the head groups are more directed towards the water layer, probably as a consequence of the interruption in the head group interactions caused on placing a foreign body in the hydrogen bonding network that holds the interface together. The helix also causes the two acyl chains to lose their heterogeneous fluidity, as the Sn2 acyl chain experiences reduced mobility and takes up a more parallel alignment with the Sn1 chain. The structural internal data (which describes the motion of the lipid molecules in specific non-bond angles and lengths) and the torsion standard deviations indicate that a reduction in the glycerol motion exists and the SOP data indicates that the Sn2 chain loses gauche population. Thus, the structural results show that the observed loss of Sn2 chain motion results from a combination of reduced methylene group isomerisation and reduced flexibility in the glycerol group.

The P-P distance for the single helix simulations was ≈37.5 Å and thus is 1.5 Å above the value for the Hex simulation, indicating altered lipid structure and gives a value as close to the Sq simulation as to the Hex simulation. The water RDF and the DPF plots indicate that the water structure and the distribution of the important groups of atoms across the bilayer is similar to that of the Hex simulation.

It is commonly accepted that when a helix is incorporated into a fluid bilayer system and there is a disparity between the hydrophobic width of the bilayer (approximately the acyl chain region) and the peptide length, the bilayer adjusts to accommodate the peptide in order to optimise non-bond interactions.[2-4] Usually if the peptide spans the hydrophobic width then the more flexible bilayer structure reduces acyl chain motion and therefore increases the bilayer thickness, thus reducing the disparity.[3,4]

As mentioned before, the helices used in this simulation were designed so that the termini penetrated the surface of the bilayer in order to investigate the interaction of the head groups with the mutated residue and also to replicate the protein helix situation. Therefore, the length of the helix used in the simulations automatically created a disparity between itself and the bilayer interior. The reaction of the bilayer to this disparity, as discussed above, can be observed to reduce the mobility of the Sn2 chains, increasing the correspondence of the acyl chains motion. Thus, it is clear that the lipid bilayers became more structured in order to accommodate the helices.

Shen et al observed that a 32-mer alanine peptide affected the SOP and acyl chain motion of a DMPC bilayer but did not significantly affect other bilayer properties such as the diffusion rate. They also observed a 1 Å change in the P-P distance.[5]

Another explanation of this effect could be that the helix causes a discontinuity in any common or semi co-operative lipid motion. This common motion could be a consequence of the initial geometry of the model bilayer system. In the results and discussion on the relative merits of the cubic and hexagonal pact bilayer it was concluded that using a cubic geometry produced a co-operative tilting motion which was not a property observed in L_α phase bilayers. This artefact of the building process resulted in the use of the hexagonal geometry as a starting point for these simulations. The bilayer SOP profiles (see Supplementary Work for detail) tend to suggest that either a residual amount of co-operative motion remains or a common mode exists that is interrupted by the helix. Also, the high torsion barrier for the methylenes could be contributing to the lower mobility of the chains, but as these observations are relative to other simulation the loss in torsion fluidity will have other contributing factors.

In previous[6] work we found that motion in the glycerol was the lowest frequency mode of motion in an isolated DPPC molecule. Using an approximation of normal mode analysis for the coordinates of a MD trajectory we calculated that the scissoring style motion was the major component in the lowest energy mode.[6]

In conclusion, the bilayer-peptide systems are stable and reproduce certain typical bilayer features whilst also exhibiting characteristics specific to this type of system, in agreement with available data.[5]

6.6. Helix Structural Data

In previous work[7] we observed that the incorporation of helical peptides into a certain lipid bilayer systems caused the helix to develope a bend and the helix became less stable as a consequence of the lipid motion. This was postulated to occur because of artefacts imparted by the starting structure of the lipids. The tilting of the cubic packed lipids has been tested in this study and confirmed to occur as a result of the cubic lattice and the resultant $L_{\beta}^{'}$ phase style bilayer structure, whereas the hexagonal geometry reproduces a good L_{α} phase.

The helix structural data indicates that no bend is observed in any of the seven single helix incorporated simulations and they all have stable, cylindrical and typical α_R structure. The only real exception to this result is the Hex_Asp system where the formal charge of the aspartate side chain appears to destabilise the helix core residues and alter the dynamics of the terminal residues. In the other simulations residues 4-29 exhibit good α_R character except in the Hex_Val and Hex_Ala1 systems which have a slightly reduced core residue stability, with residues 25-29 exhibiting reduced α_R structure

The helices all appear to have significant C-terminal fraying or a more defined tendency to fray at the C-terminus than the N-terminus, except for the Hex_Asp system simulation where it appears to have a more unstable and frayed N-terminus. Therefore, it exhibits contrary terminal dynamics to the neutral aspartic acid. Thus, systems Hex_Ala1, Hex_Asp-, Hex_Arg, Hex_Thr and Hex_Val exhibit highly mobile and/or formally frayed C-terminal structure over and above that of the N-terminal.

The Hex_Phe system simulation indicates that it too has highly mobile C-terminal residues, but it has an equal amount of reduced α helical geometry at the N-terminus as at the C-terminus. Whilst, the Hex_Phe system data indicates similar terminal structural behaviour, the C-terminus still has a slightly higher degree of mobility than the N-terminus. Thus, it can be argued that the reduced fraying of the C-terminal

is due to the Phe ring of residue 29 protecting the exposed terminal hydrogen bonds from the polar waters atoms.

The mutation site at residue 17 has been observed to affect the structural stability of the central residues. This slightly altered structure was observed in all systems except the Hex_Phe system where no change in the mutation site associated residues is indicated. The hydrogen bond lengths for residues 13, 14 and 17, for the 6 effected systems, are slightly higher than the adjacent hydrogen bond lengths, indicating that back-bone of the 17th residue is less ideal and the associated hydrogen bond is weakened.

The stability of the core residues in the 7 systems can be ranked according to the results described above:

Hex_Ala1, Hex_Asp, Hex_Arg, Hex_Thr, Hex_Phe > Hex_Val > Hex_Asp-

This ranking of the helices highlights that the mutation of the centrally located residue has a small effect on the stability of the helix in the lipid bilayer. Only the formal charge of the aspartate side chain has a significant affect on the helix core residues. This is possibly due to two reasons. Firstly, the charged COO^- group of the Asp side chain could form a hydrogen bond back to the backbone amide, interrupting the backbone hydrogen bonding or secondly the negative charge could either be repelling or attracting backbone atoms and causing them to lose helical geometry. The reasons for this and other interesting results will be investigated further in Chapter 8.

The interaction of the helix dipole with charged residues at the termini of helices is well known to affect stability, with negative charges stabilising the C-terminus and positive charges stabilising the N-terminus. The structural data does not confirm this to be the case for membrane proteins. Probably as a result of the charged side chains preferring to interact with the head group waters.

The increased helical core mobility and lower ideality for the Hex_Val system is a little surprising as the valine side chain is hydrophobic and would be expected to be well suited to being placed in the lipid bilayer core region. The cause for this reduced ideality could be the fact that the side chain is hydrophobic, which allows it to achieve a better energetic stability by interacting with acyl chains of the lipid molecules, which will compensate for the loss of helicity..

The mobile terminal residues were expected to populate the 3_{10} geometry as it has been observed to be an intermediate in the unfolding of helices.[8,9] Shen et al [5] observed 3_{10} structure in the fraying termini of a polyalanine helix in a bilayer MD study. The single helix simulations do not show any 3_{10} geometry even in the most frayed residues. Although the hydrogen bonding data did not indicate the presence of 3_{10} structure the dihedral torsion angles did indicate that another weak geometry was present, with a more negative ϕ angle than the 3_{10} geometry. The hydrogen bond data did indicate that at specific residues when ideal alpha helicity was lost another stable conformation was populated. This conformation is in the general helical region of the $\phi\psi$ map. A possible explanation of the occurrence of this non-classical geometry is that potentials used in the V.F.F. favours slightly more negative dihedral angles than those used in the initial helix.

Thus, all the data confirms that the mutated helices have all incorporated well into the bilayer with no significant artefacts introduced to either the bilayer or peptide structures. Hence, the extension of the analysis to the energetics in Chapter 8 is valid.

References

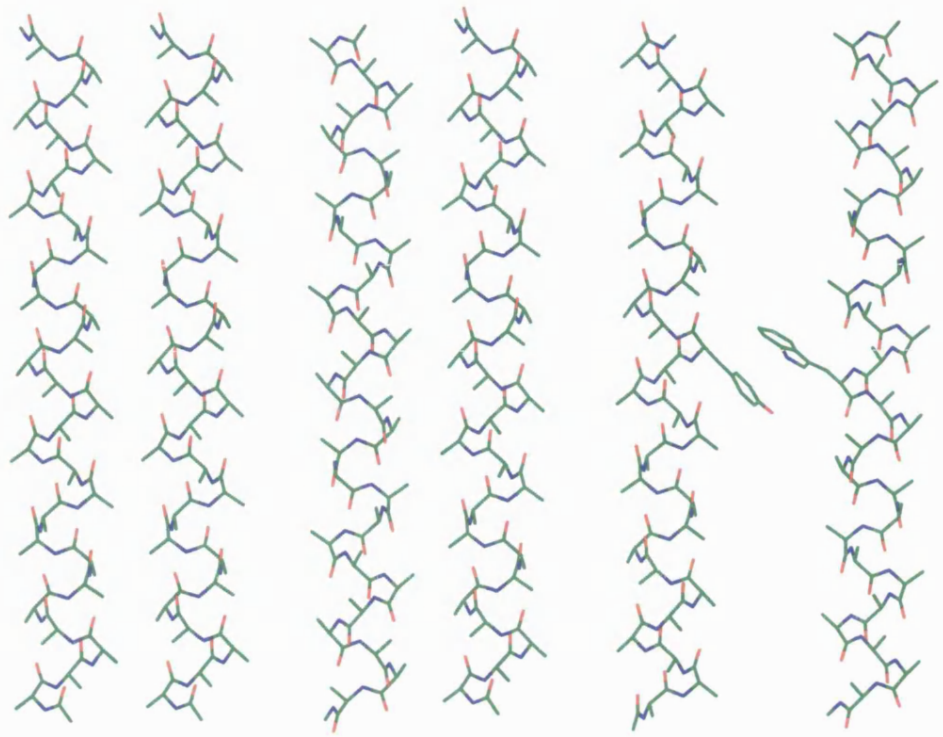
1. Biosym. Technologies Inc. San Diego, CA, USA
2. F. A. Nezil and M. Bloom, *Biophysical Journal*, **61**, 1176-1183 (1992).
3. Y. P. Zhang, R. N. A. H. Lewis, G. D. Henry, B. D. Sykes, R. S. Hodges, and R. N. McElhaney, *Biochemistry*, **34**, (7), 2348-2361 (1995).
4. J.A. Killian, *Biochimica et Biophysica Acta-Reviews on Biomembranes*, **1376**, 401-416 (1998).
5. L.Y. Shen, D. Bassolino, and T. Stouch, *Biophysical Journal*, **73**, 3-20 (1997).
6. G. E. Adams. Final Year Report M.G.U., Bath, 1993
7. G. E. Adams. M.Phil. Thesis M.G.U., Bath, 1995
8. W. R. Fiori, S. M. Miick, and G. L. Millhauser, *Biochemistry*, **32**, 11957-11962 (1993).
9. W. S. Young and C. L. Brooks, *Journal Molecular Biology*, **259**, 560-572 (1996).

7. Helix Pair Incorporated Bilayer Simulations

7.1. Building the Helix Pairs

The 30-mer polyalanine blocked peptide used in the single helix simulation, Hex_Ala1, was used to create the helix pair systems. A pair of the polyalanine α_R helices were placed in a mutual parallel orientation at a separation just outside their Van der Waal (VdW) surface to ensure good packing. The anti-parallel helix pair was constructed by rotating through 180° one of the helices and the helices were placed so that their separation again allowed good VdW packing.

A third helix pair system was constructed to specifically reproduced the Tyrosine-Tryptophan residues of helices A and B found in the transmembrane helical section of Bacteriorhodopsin in order to probe the interactions of this arrangement when in a membrane environment. The anti-parallel helix pair was used as the host for the residues, with the 17th residue of helix 1 and 18th of helix 2 being mutated to a Tryptophan and Tyrosine, respectively. The separation of the two helices was increased in order to accommodate the bulky side chains of the mutated residues. The two aromatic rings were aligned as close as possible to that observed in the Bacteriorhodopsin native structure, given the geometric constraints of the bilayer host system. Fig. 7.1 illustrates the helix pairs used in the three helix pair simulations.



a b c

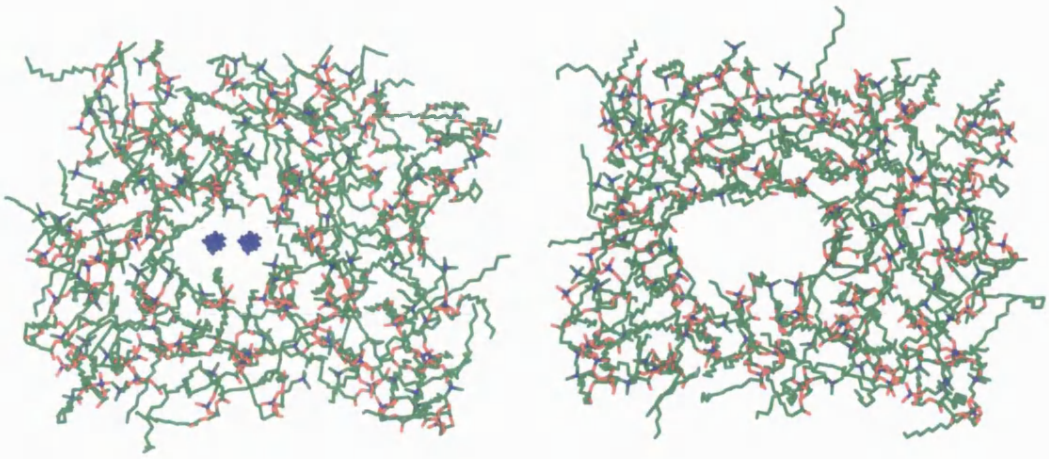
Fig. 7.1 - The Three Helix Pair Systems (Hex_Ala2p-a,Hex_Ala2a-b,Hex_Trp_Tyr-c)

7.2. Building The Bilayer Cavity

The hydrated DMPC bilayer structure resulting from the hexagonal packed geometry simulation (Hex) was used to construct the host bilayer cavity.

In order to produce a cavity in the centre of the bilayer 2 pairs of lipids were removed and as with the single helix simulation, a pair of fake fixed geometry atom spirals were placed in the space produced. The VdW surface of the spiral pair was gradually increased, over a 20 ps normal MD simulation, in order to repel out of the cavity the fluid acyl chains of the lipid molecules adjacent to those removed. The spiral's VdW radius initially was set to that of a spiral of hydrogen atoms and was

increased to near that of the helix pairs. The bilayer before and after the cavity production procedure are illustrated in Fig. 7.2.



a b

Fig. 7.2 - Helix Pair Cavity (Before-a,After-b)

As with the cavity production in the single helix systems, it was important that the surface area occupied by the spiral was monitored so that it did not exceed that of the lipid pairs removed, ≈130 $Å^2$. After the first 10 ps of this simulation the surface area of the spiral pair equaled that of the 2 lipid pairs removed, but its was also discovered that at this point the VdW volume of the spiral pair was below that required. The expansion process was continued for a further 5 ps until it was observed that the acyl chains had moved sufficiently to allow the clean insertion of the helix pair and the simulation was stopped at this point. The dimensions of the unit cell were adjusted during the 11-15 ps phase of this simulation in order to maintain the lipid bilayer area per head group (SA) to that of the parent Hex simulation.

The non-bond energy (Fig. 7.4) was monitored during this process to follow the effect that the repulsion of the acyl chains had on the system. The non-bond energy, whilst rising a little more than the corresponding single spiral simulation, was well within that expected for such a process. Once the cavity was found to be at the required dimensions the simulation was carried on for another 30 ps to allow the

system to settle down before the helical peptides were introduced.

During the above process, the inter-bilayer water layer was retained so that the bilayer structure remained as near as possible to its original Hex simulation structure.

After the cavity production procedure, the three α helical pairs were placed in the bilayer with their long axis along the bilayer normal and with the residue 17 of helix 1 in the centre of the bilayer acyl chain region.

Hex_Ala2p - Parallel polyalanine helix pair bilayer simulation

Hex_Ala2a - Anti-Parallel polyalanine helix pair bilayer simulation

Hex_Trp_Tyr - Anti-Parallel polyalanine helix pair, with Tryptophan and Tyrosine residues mutated, bilayer simulation

7.3. Simulation Details

7.3.1. Minimisation

The three bilayer systems (Hex_Ala2p, Hex_Ala2a, Hex_Trp_Tyr) were minimised to remove clashes introduced during the building process. During the minimisation the coordinates of the helices were fixed in space, to their original values, in order that the hydrated lipids did not adversely affect the helical structure.

Periodic boundary conditions were applied and the unit cell vectors were retained from the cavity building procedure.

The systems each had a number of clashes either specific to bilayer/water interactions or to the Trp and Tyr residues. The closely related Hex_Ala2p and Hex_Ala2a systems exhibited clashes between the lipid head groups and between lipid head groups and water molecules that resulted from the Unit Cell Vector (UCV) scaling procedure. The head group clashes were removed through applying the steepest decents minimiser for 70 steps to the clashing molecules only and the whole system was minimised for a further 10 steps of steepest descents minimisation to remove

residual clashing interactions. The Hex_Trp_Tyr system required careful minimisation to remove the clashes whilst not effecting the equilibrated lipid molecules. The steepest decents minimiser was again only applied to the clashing lipid molecules and the solvating water molecules for 200 steps to remove the strain induced during building. Following this minimisation the Trp and Tyr residues only were minimised for 100 steps of steepest decents minimisation and the whole system was finally minimised for another 100 steps of steepest decents minimisation to again remove residual clashes.

7.3.2. Molecular Dynamics

The 3 minimised systems were then run under the V.F.F. for 100 ps using a 1 fs time step and with the recording of the velocities and the coordinates every 20 fs. A truncated non-bond cut-off was used for the Lennard-Jones and coulombic interactions at 16 Å, with a 18 Å cut off being applied to the neighbour list generation. Period boundary conditions were applied and the unit cell dimensions were retained from the building process. NVT conditions were also applied as with the Hex and Sq simulations.

The initial system temperature for the random Boltzmann distribution of velocities was set at 200 K and raised over the first 5 ps to the simulation temperature of 311 K. A multiple temperature bath was used to maintain the simulation temperature to 311 K, with an individual bath being applied to the helix pair, the bilayer and the inter-bilayer water layer. The helix pair coordinates were restrained to their starting values during the same 5 ps period in order to allow the system to relax whilst not influencing the helix structure.

The velocities were reassigned every 1 ps during the initial 5 ps period in order to improve the randomisation and equilibration of the initial structure.

7.4. Bilayer System Results

An extensive analysis of the bilayer structure has been carried for the three helix pair simulations. This analysis was identical to that for the pure bilayer simulations reported in Chapter 5. The structure of the peptides in the helix pair simulations was analysed and gives an insight into the dynamics of such helical arrangements when incorporated into a bilayer environment.

The thermodynamic and temperature results for each of the three helix pair simulations are included in this chapter. The bilayer and helix pair structural results are referred to in the discussion and the important results are highlighted, with the other results included in Supplementary Work, Section II.

7.4.1. Average Temperature - Hex_Ala2p

Fig. 7.3 contains the group temperatures across the simulation for the parallel polyalanine helix pair (Hex_Ala2p) system. The different sub system groups were calculated separately from a sum of the individual atomic velocities. Table 7.1 contains the simulation averages and the standard deviations (S.D.).

Table 7.1

Hex_Ala2p Average Temperatures, K		
	Temp	S.D.
System	312.2	1.6
PC	314.1	6.8
GL	313.5	9.1
Sn1	308.3	4.5
Sn2	311.2	4.0
Water	313.9	2.6
HC	309.7	2.7
Top Mon	310.8	3.2
Bot Mon	311.6	3.2

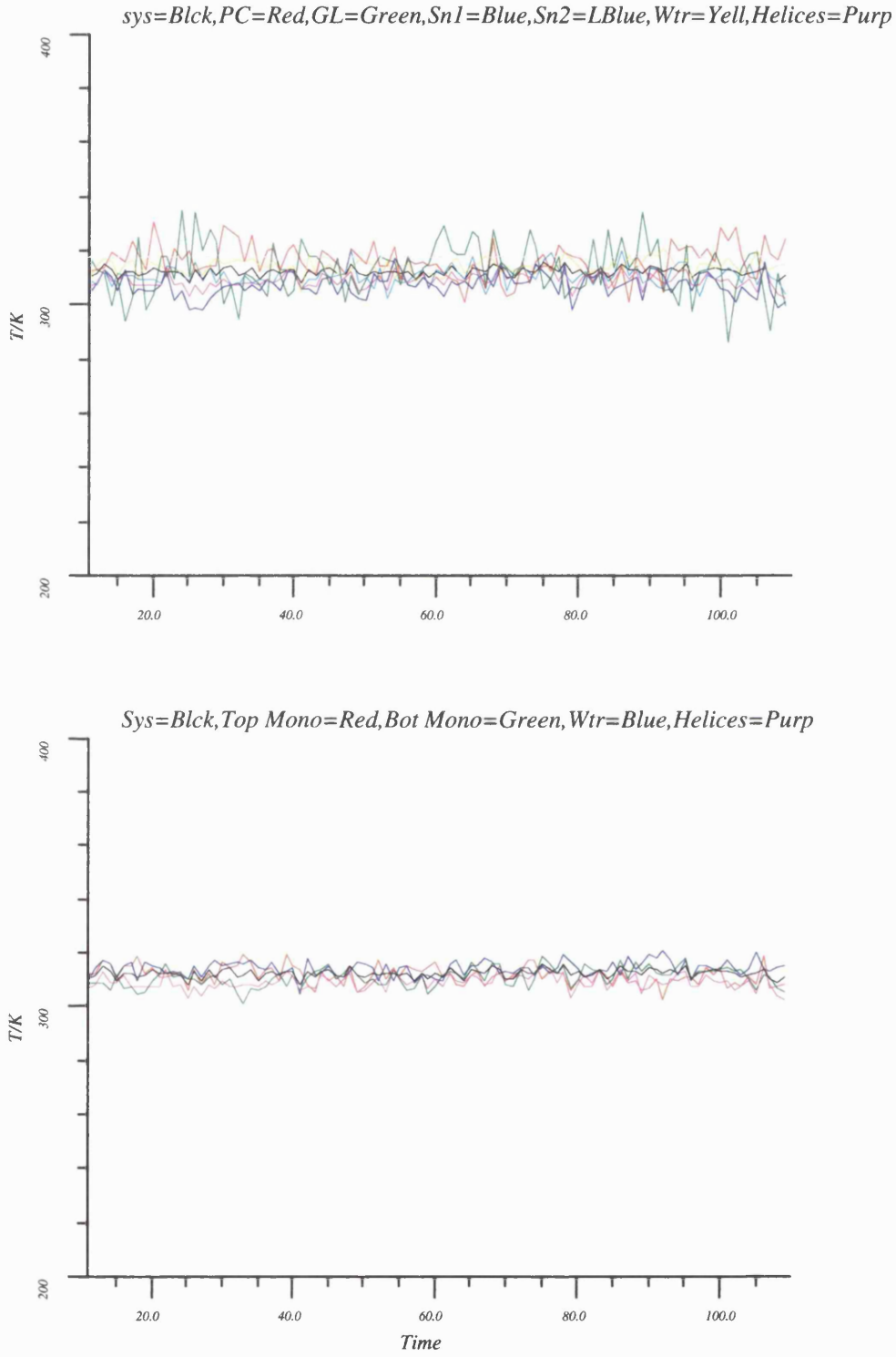


Fig 7.3 - Hex_Ala2p Average Temperature

As can be seen from Fig. 7.3 the average temperatures for the different parts of the system are very mobile and indicate that different sub-sections of the system possess slightly different average temperatures. The more polar groups of atoms appear to produce peaks above the system temperature bath value of 311 K, whilst the less polar groups of atoms fluctuate less. The simulation averages in Table 7.1 clearly show similar behaviour to that exhibited by both the Hex and Sq systems, with the group temperatures influenced by the cut-off temperature effect, as discussed previously in Chapter 4.

7.4.2. Thermodynamic Data - Hex_Ala2p

The general system conditions and the thermodynamic data for the Hex_Ala2p system are illustrated in Fig. 7.4.

The unit cell vectors (UCV) are observed to be constant throughout the simulation as would be expected for a system being constrained under NVT conditions.

The warm-up phase of the simulation, where the temperature of the system was gradually increased to the simulation temperature of 311 K, can be clearly seen in the kinetic, potential and total energies, which increase in value over the first 5 ps. After this initial period all three energies are equilibrated.

The same 5 ps warm-up phase can be observed in the bond and theta energies. This is as would be expected for these high frequency degrees of freedom but the phi energy, which describes the torsional motion of the molecules, exhibits a different behaviour to the bond and theta energy. The phi energy appears to undergo a transition to a higher value and after 5 ps it begins to relax to a lower energy. After 15 ps of the simulation it is equilibrated. The relaxation observed in the phi energy results from an equilibration of the two parallel helices interactions and the lipid molecules adjusting to the helix pair.

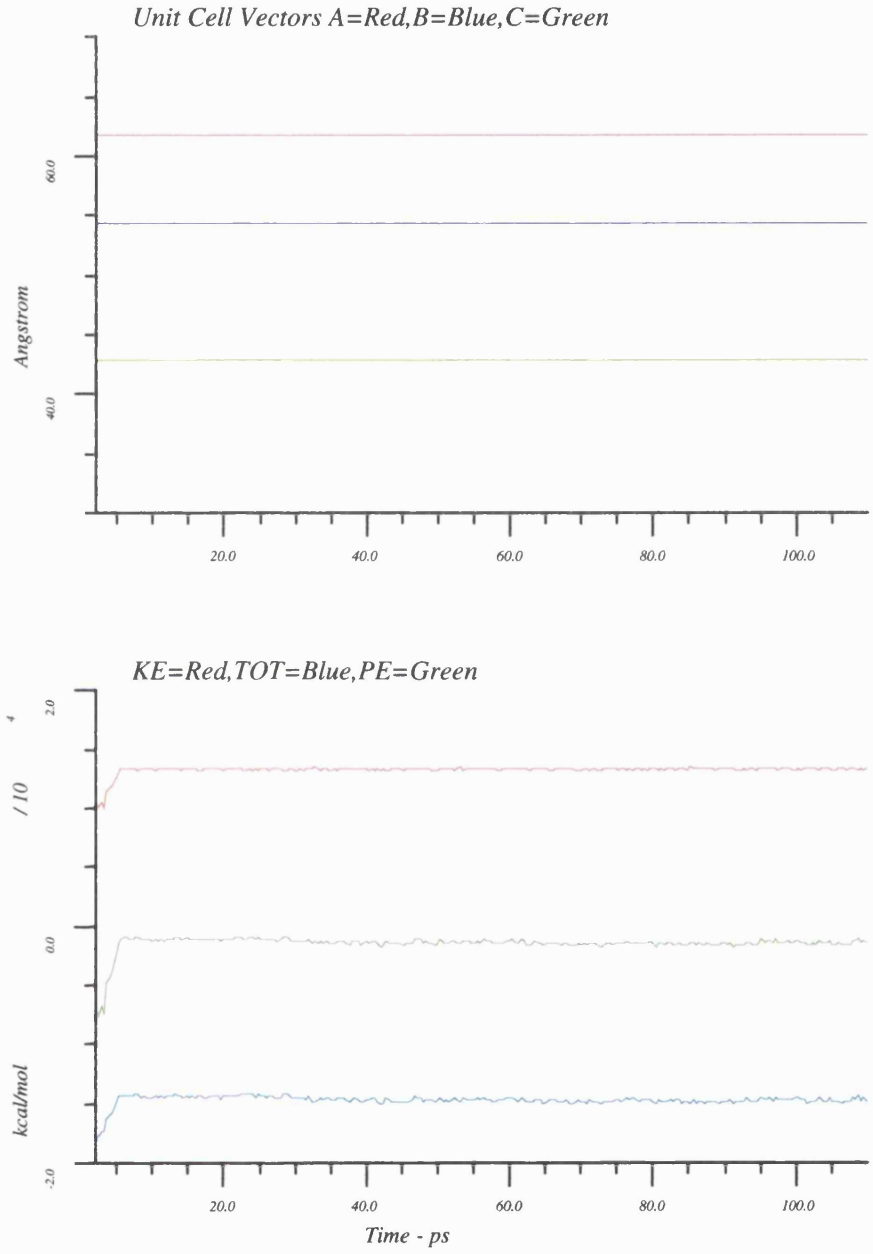


Fig.7.4a - Hex_Ala2p Thermodynamic Data

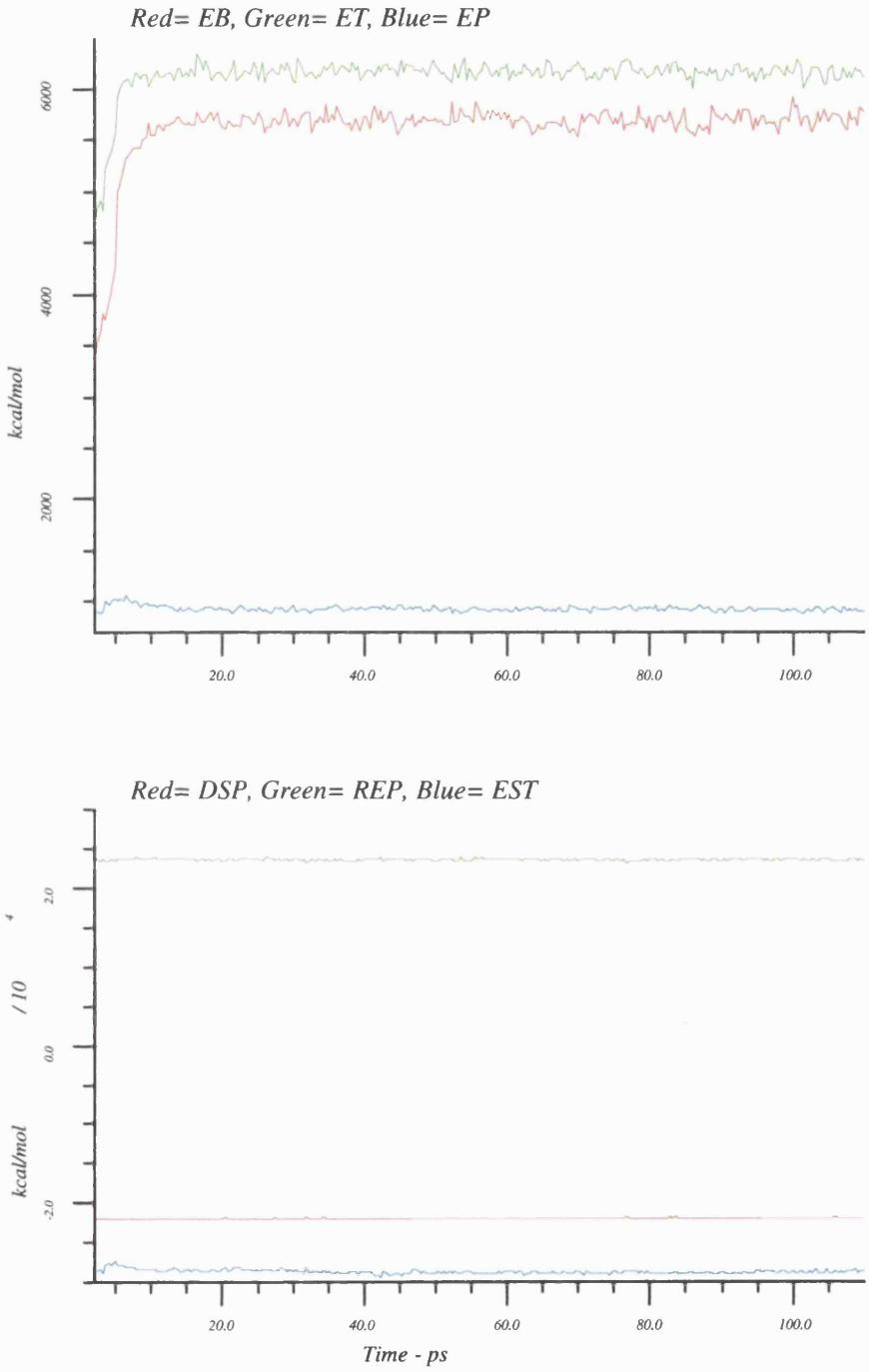


Fig.7.4b - Hex_Ala2p Thermodynamic Data

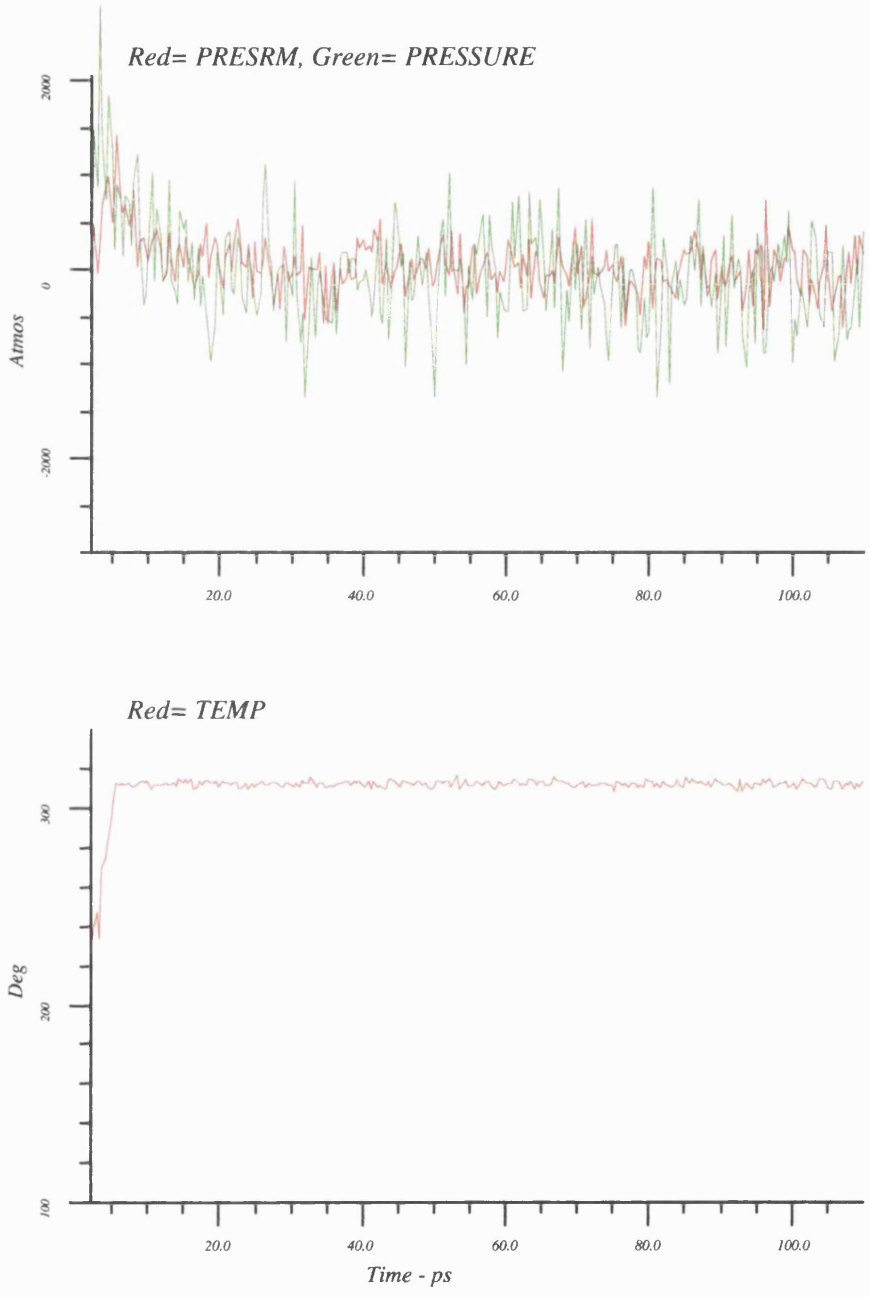


Fig.7.4c - Hex_Ala2p Thermodynamic Data

The non-bond energy components in Fig. 7.4 indicate that the Van der Waals component energies (repulsive and dispersive) are not sensitive to the increased thermal energy, whilst the electrostatic energy shows a degree of change early in the simulation. Initially the electrostatic energy increases slightly in response to the increasing kinetic energy of the system but after 5 ps it starts to relax back to its original value, as observed with the torsional energy. This could result from the increasing kinetic energy affecting the structural integrity of the helices, in which much of the stability is derived from electrostatic interactions.

The observed changes in the electrostatic energy could also be the result of thermally enduced disruption in the well ordered head group region and indicate that the helix is becoming distorted from its initial structure. At the end of the warming up phase the system is equilibrated.

The thermodynamic temperature of the system is observed to be equilibrated throughout the simulation and as with the kinetic energy the warming up phase of the simulation is observed also.

The system and molecular pressures show interesting behaviour for the initial period of the simulation, with the warm-up phase appearing to induce a higher pressure which relaxes as the simulation progresses and is equilibrated by 10-15 ps. The pressure is fluctuating near to 0 atmospheres. Clearly the warming-up of the Hex_Ala2p system induces a change in the system dynamics that is either lost or is equilibrated at 15 ps.

Thus, the thermodynamics data indicates that the system is equilibrated with the warm-up phase of the simulation causing small but observable changes in the torsion energy, the electrostatic energy and the pressure. This indicates a conformational change in the system that is triggered by the increase in thermal energy early in the simulation and is likely to have occurred in the helices as such a structural alteration would affect the electrostatic energy and the pressure of the system.

7.4.3. Average Temperature - Hex_Ala2a

Table 7.2

Hex_Ala2a, Average Temperatures, K		
	Temp	S.D.
System	312.1	1.6
PC	315.6	5.8
GL	314.8	9.1
Sn1	309.8	4.3
Sn2	308.5	4.4
Water	313.8	2.5
HC	309.2	2.7
Top Mon	311.3	3.4
Bot Mon	311.1	3.5

Fig. 7.6 contains the group temperatures across the simulation for the Hex_Ala2a system. The different sub system groups were separately calculated from a sum of the individual component atom velocities. Table 7.2 contains the simulation average values and the standard deviations of the averages (S.D.).

The average temperatures indicate that, as with the previous simulation, multi temperature bath fails to fully remove the cut-off induced temperature effect. As described before the heterogeneous nature of the bilayer system means that several temperature baths need to be used if a near exact temperatures are required, but using 3 individual baths is sufficient in this case.

Thus, the system temperature places the DMPC bilayer in the correct region of the phase map for the L_α phase.

7.4.4. Thermodynamic Data - Hex_Ala2a

The general system conditions and the thermodynamic data for the anti-parallel polyalanine helix pair (Hex_Ala2a) system are illustrated in Fig. 7.5.

The unit cell vectors (UCV) are constant throughout the simulation as would be expected for a simulation performed under NVT conditions. The kinetic, potential and total energies are stable from 5 ps to the end of the simulation and are close to the values observed for the Hex system and the single helix incorporated systems. The initial warm-up phase of the simulation can clearly be seen over the first 5 ps where the temperature was gradually increased.

The bond, theta (angle) and phi (torsion angle) energies, which describe the internals of the lipid molecules, indicate that the system is stable throughout the simulation. During the initial 10 ps of the trajectory, the higher energy theta and bond energies exhibit an increasing value as the thermal energy increases, whilst the torsion energy appears to undergo changes. The theta and bond energies appear to slightly reduce as they equilibrate over the next 10 ps and then fluctuate around their average values.

The non-bond, electrostatic and Van der Waal's (VdW) energies (repulsive and dispersive), also show differential profiles during the warm-up phase and the remainder of the simulation. The VdW's energies appear to undergo only a very slight reaction to the warm up phase and then remain essentially constant through the remainder of the simulation. The electrostatic energy of the system appears to undergo a transition during the warm-up phase. This could be a result of conformational changes in the helix molecules or in the head group region as the waters and PC atoms become thermally active as both these factors effect the electrostatic interaction energy.

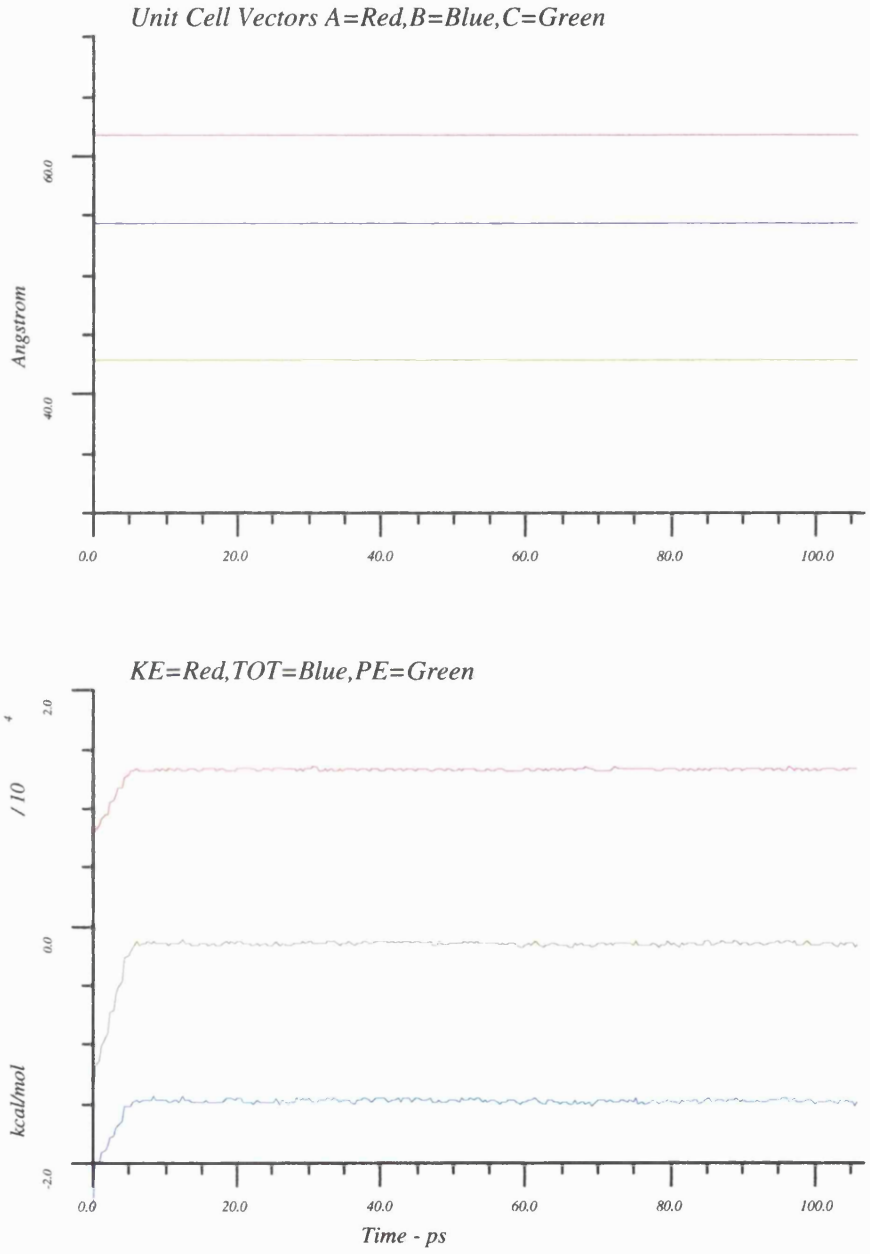


Fig. 7.5a - Hex_Ala2a Thermodynamic Data

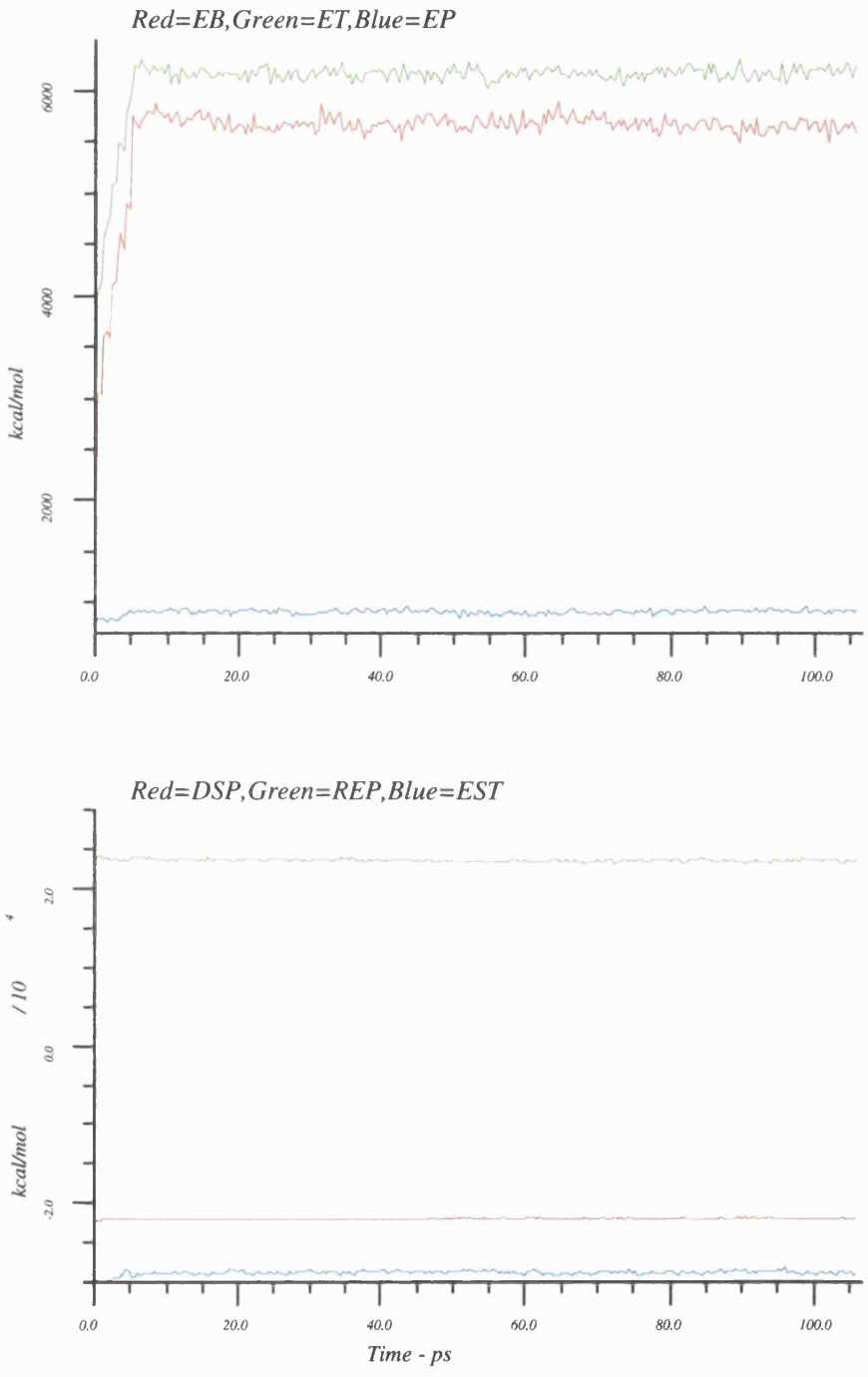


Fig. 7.5b - Hex_Ala2a Thermodynamic Data

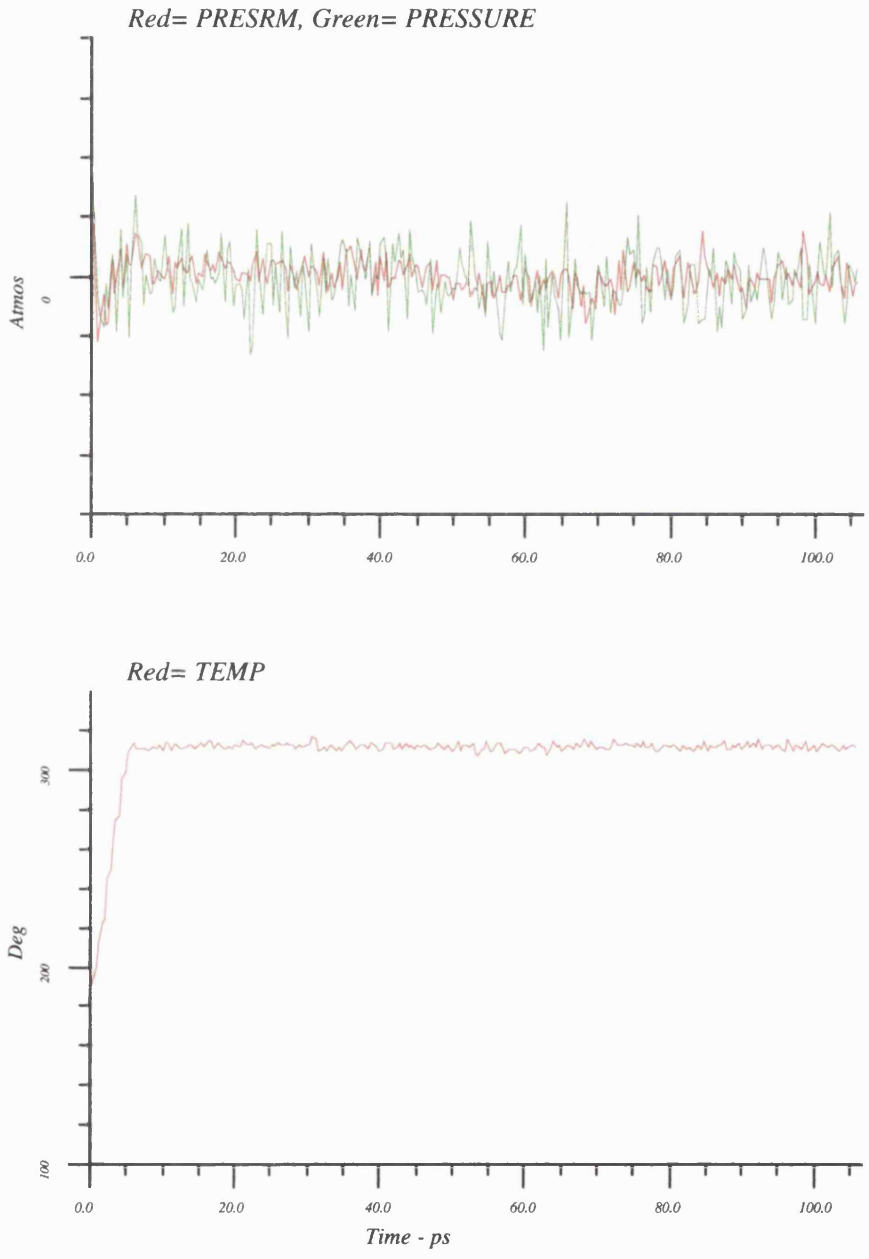


Fig. 7.5c - Hex_Ala2a Thermodynamic Data

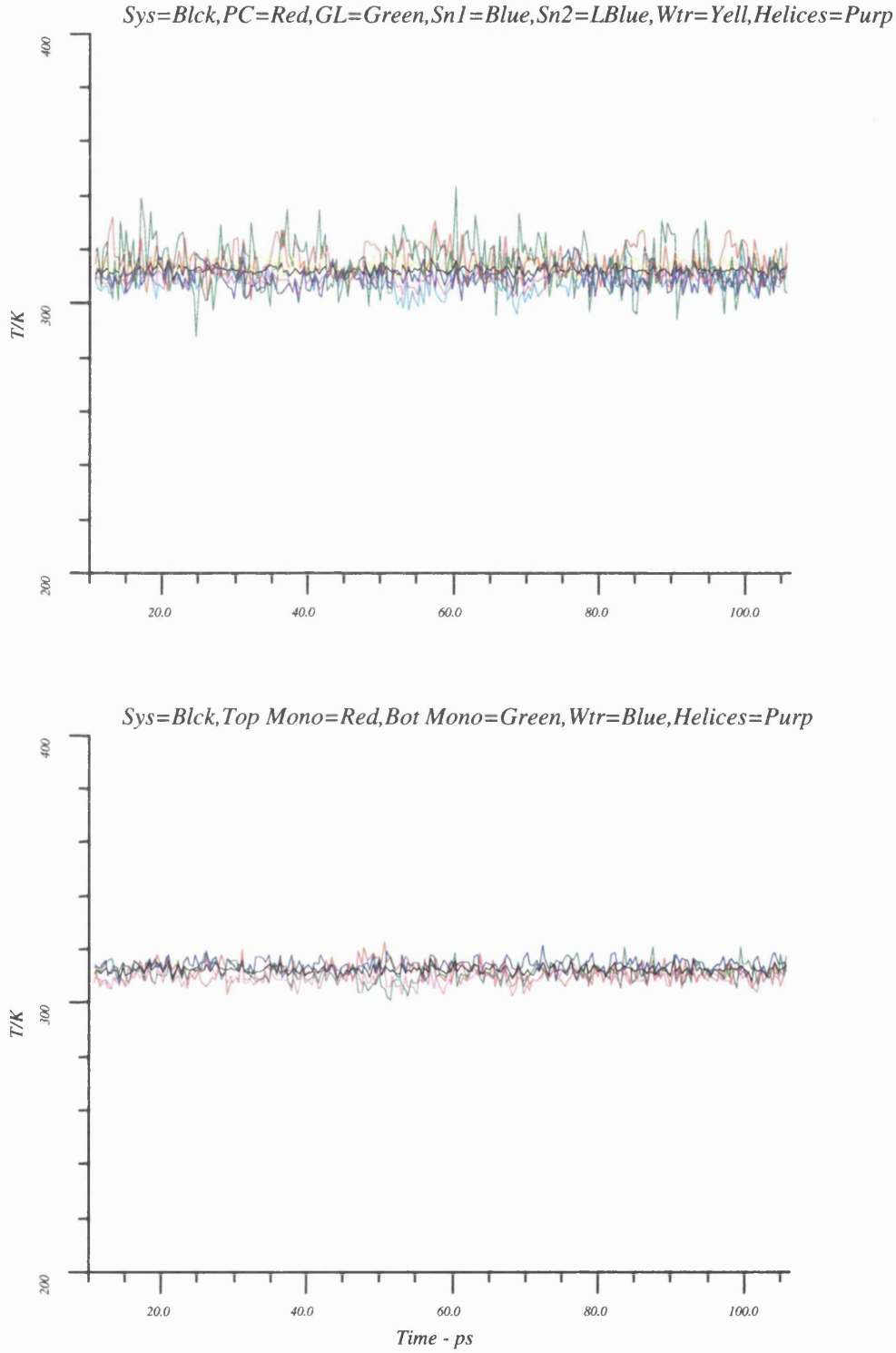


Fig. 7.6 - Hex_Ala2a Average Temperature

The molecular and system pressure both indicate that the system is equilibrated around 0 atmospheres. The use of a NVT conditions appears to have no adverse affect on the pressure of the system.

The thermodynamic temperature is stable as would be expected for a system that has a temperature bath applied to it.

Thus, the data indicates that the system is thermodynamically equilibrated with only slight changes occurring as a result of the warming up of the system and the helical restraints during the first 5 ps.

7.4.5. Average Temperature - Hex_Trp_Tyr

Fig. 7.8 contains the group temperatures across the simulation for the Hex_Trp_Tyr system. The different sub system groups were separately calculated from a sum of the individual component atom velocities. Table 7.3 contains the simulation average values and the standard deviations (S.D.).

Table 7.3

Hex_Trp_Tyr, Average Temperatures, K		
	Temp	S.D.
System	312.2	1.5
PC	315.1	6.1
GL	311.3	9.1
Sn1	310.1	4.1
Sn2	309.4	4.1
Water	313.7	2.6
HC	309.8	2.7
Top Mon	310.7	3.2
Bot Mon	312.0	3.1

The average group temperatures for the sub sets of the system can be seen to replicate the results observed for the other simulations, with the actual temperatures

being above or below the temperature bath value. As described before, when the non-bond interactions are truncated the polar atoms warm-up and the apolar atoms cool down gradually due to the cut-off temperature effect.

This variation in atomic temperature is only a slight deviation from the 311 K required value and all the groups of atoms are above the threshold for the transition to the L_α phase for hydrated DMPC bilayers.

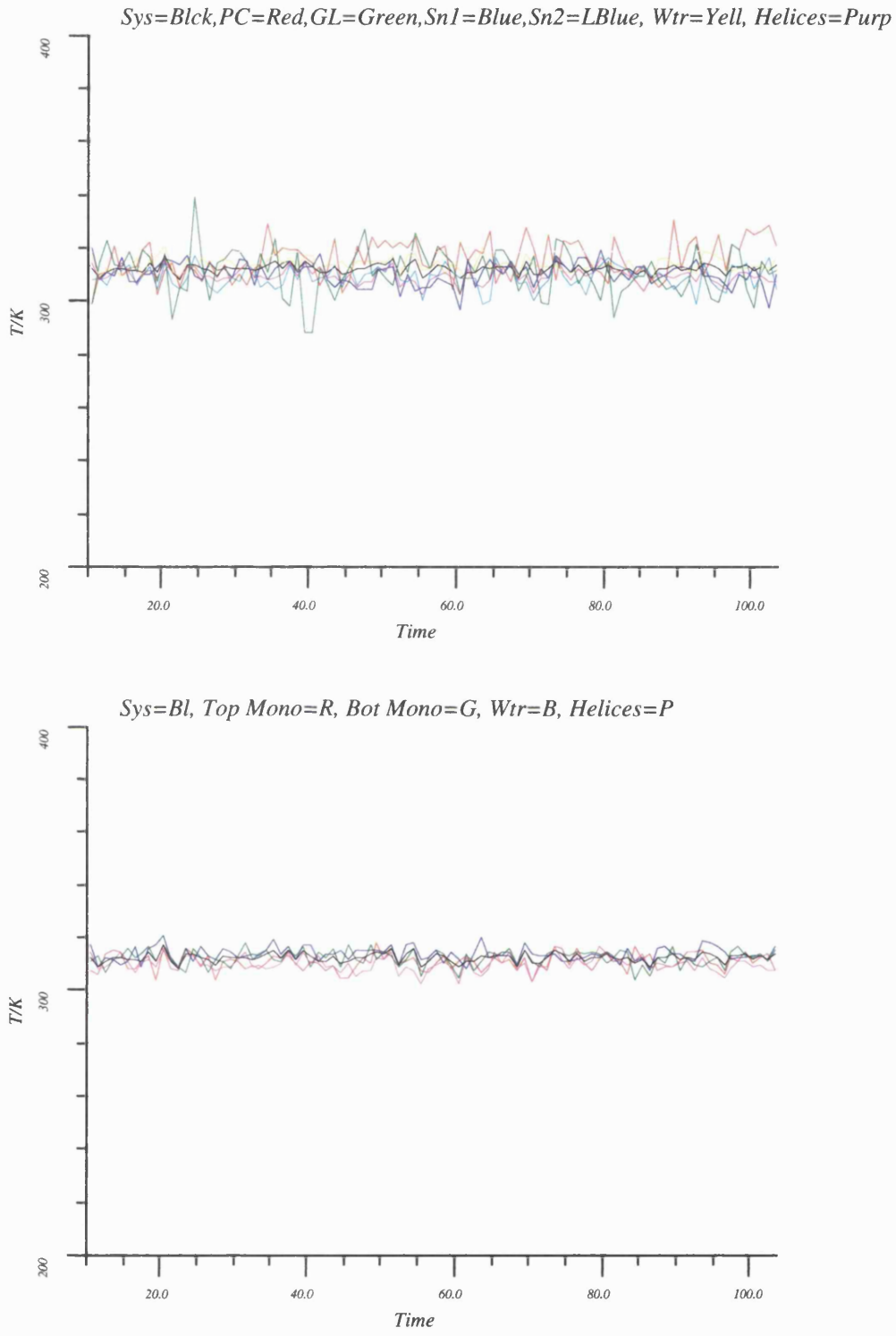


Fig. 7.8 - Hex_Trp_Tyr Average Temperature

7.4.6. Thermodynamic Data - Hex_Trp_Tyr

The general system conditions and the thermodynamic data for the anti-parallel polyalanine helix pair, with Trp/Tyr mutations, (Hex_Trp_Tyr) system are illustrated in Fig. 7.7.

The unit cell vectors (UCV), as would be expected for a simulation under NVT conditions, are constant throughout the simulation. The Kinetic, Potential and Total energies are equilibrated after the initial warm-up phase.

The bond, theta (bond angle) and phi (torsion angle) results indicate that as the thermal energy increases the bond and theta energy increases rapidly, but the phi energy is much less responsive to the increasing temperature.

The non-bond component energies (electrostatic, dispersive and repulsive) again show a response to the warm-up phase of the simulation and initial rearrangements in the structure of the system. The electrostatic energy can be seen to increase more than the other non-bond energies and probably in response to the increased motion in the head group and the loss of ideal helical structure as the helix hydrogen bonding becomes dynamic. The Van der Waal (VdW) energies react to the increasing temperature, with the repulsive energy reducing and therefore indicating slightly better packing of the system, with the dispersive energy increasing to greater extent. This small change indicates that the system relaxes by increasing some of the atom-atom distances.

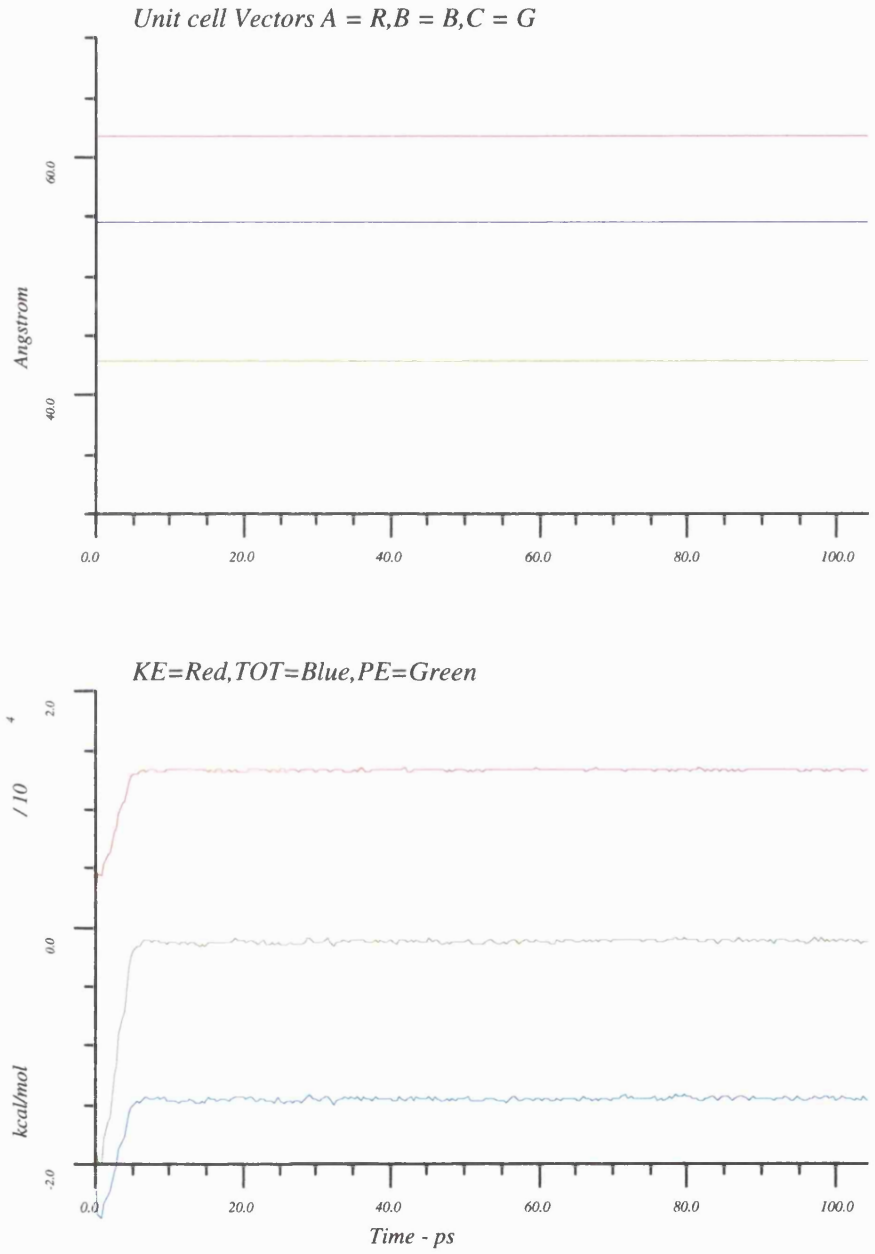


Fig. 7.7a - Hex_Trp_Tyr Thermodynamic Data

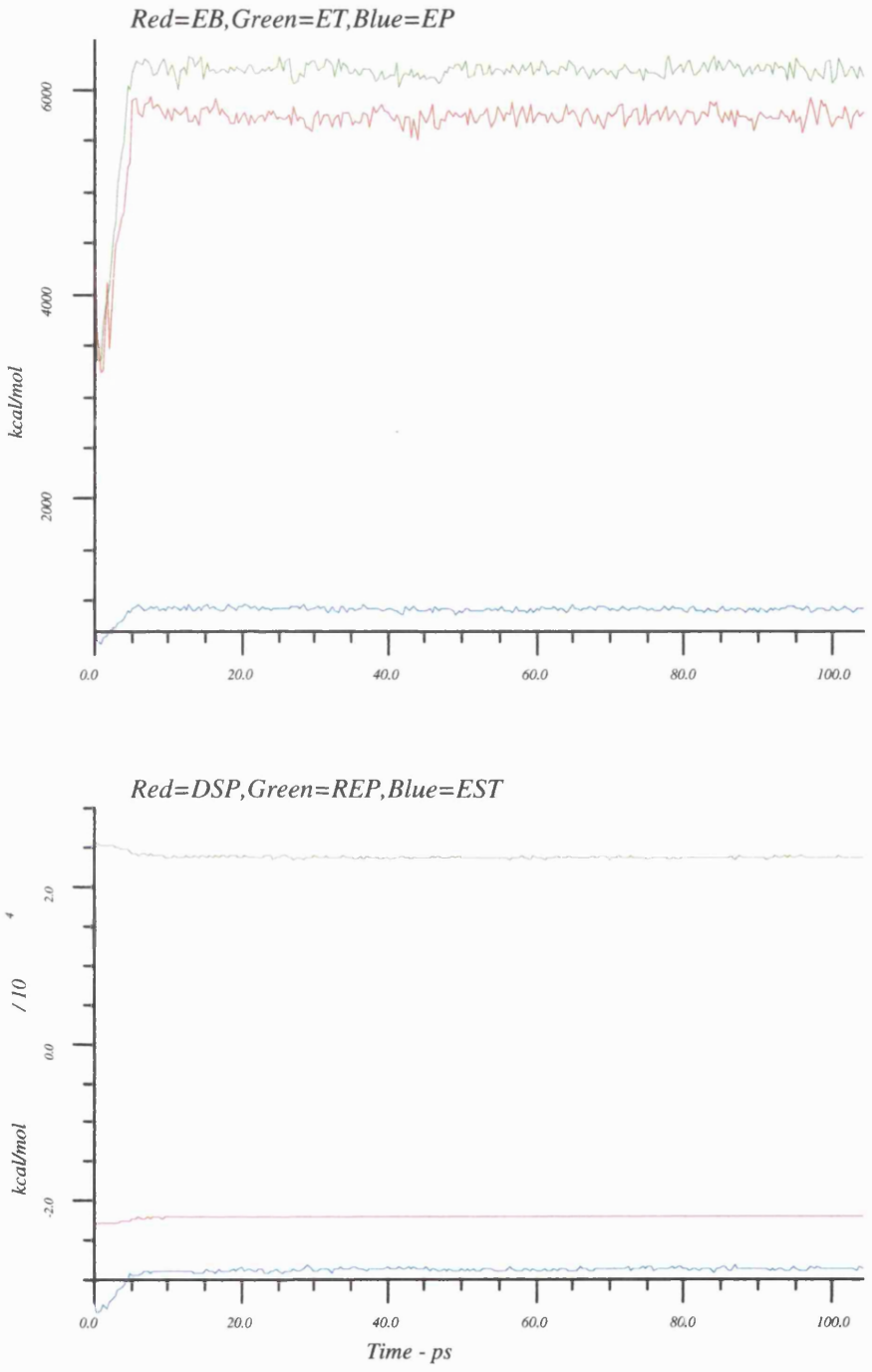


Fig. 7.7b - Hex_Trp_Tyr Thermodynamic Data

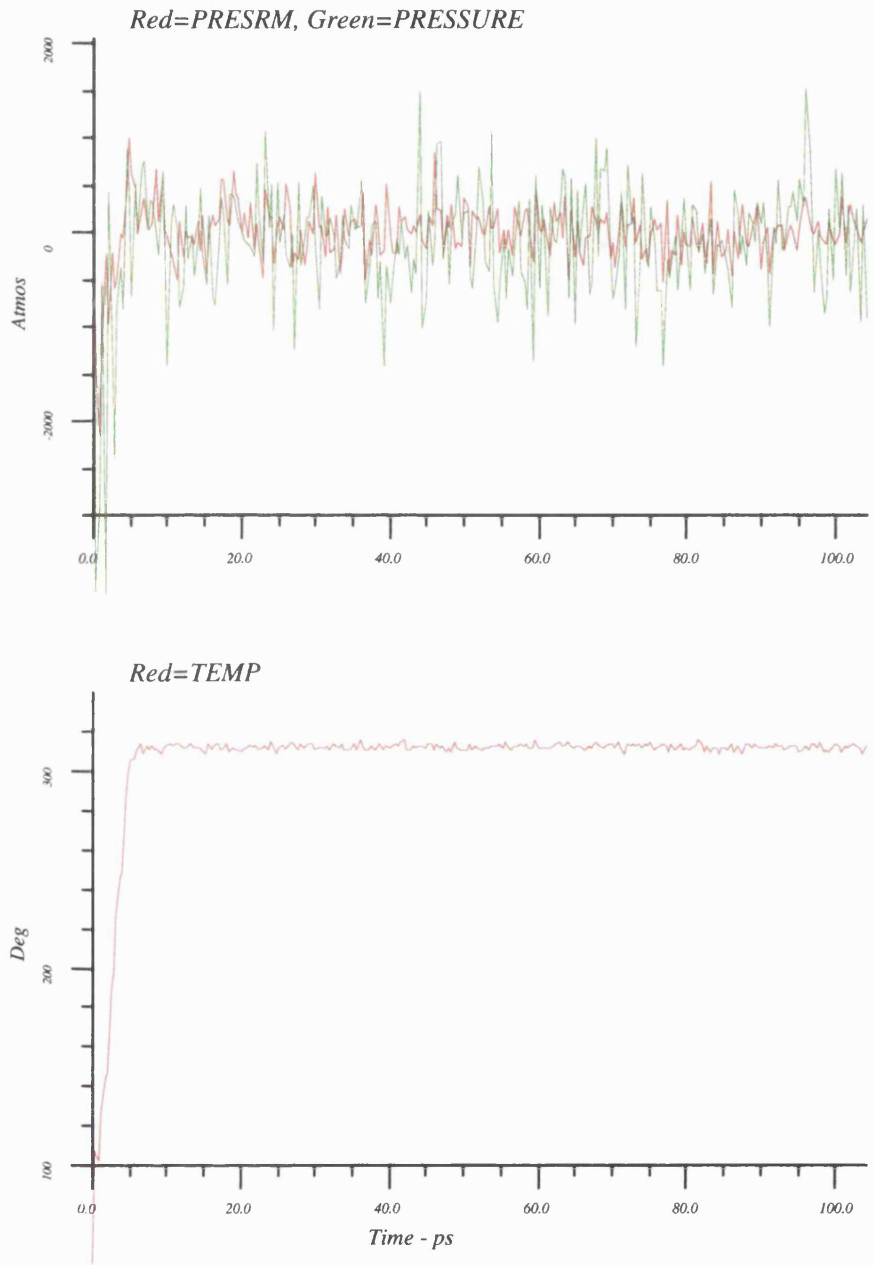


Fig. 7.7c - Hex_Trp_Tyr Thermodynamic Data

The pressures plots indicate that the warm-up phase increases the system pressure which then relaxes over 10 ps, indicating that the main section of the simulation is equilibrated.

The thermodynamic temperature is stable after the warm-up phase and indicates that the system temperature is a little above that of the applied temperature bath. This is a cut-off induced effect as discussed previously.

Thus the thermodynamic data indicate that the system is equilibrated.

7.5. Discussion on Bilayer Structural Data

The results presented for the helix pair bilayer simulations gives new insights into the interactions of a DMPC bilayer with different incorporated helix pairs and the consequences of their inclusion on the bilayer structure, as well the changes induced in the helix geometry by both the bilayer and the different helix arrangements.

The bilayer maintains most of its typical properties, such as a rough and diffuse interface region, but it also exhibits altered hydrophobic core dynamics. Interestingly, we found that the incorporation of a second helix into the system did not effect the P-P distance as the value for the three helix pair simulations was also ≈1 Å above that of the Hex bilayer system. Therefore, as with the single helix systems, the incorporation of a helix pair that has a length longer than the acyl chain region of the bilayer has an effect on the P-P distance and thus the bilayer thickness. This result again is in agreement with Shen et al.[1] As with the single helix simulations, the inclusion of a bilayer spanning molecule(s) reduces the space available to the lipid molecules and therefore, the acyl chains cannot achieve as higher a degree of fluidity as in the pure bilayer. This reduced acyl chain dynamics causes them to be more aligned and therefore, increasing the the inter-head group spacing, thus increasing the bilayer thickness. (explanation as in Chaper 6)

The major difference between the effect on acyl dynamics of the single helices and the helix pairs is that the slightly more distorted helices (explained later) in the Hex_Ala2p and Hex_Trp_Tyr simulations effected the Sn1 chain dynamics as well as the Sn2 chain. The reduced Sn1 acyl chain motion is due to just reduced trans gauche isomerisations as no other property gives significant indications of this behaviour except the segmental order parameters and the torsion data.

As well as the altered acyl chain dynamics, the incorporation of the Hex_Ala2p and Hex_Trp_Tyr helix pairs causes the head groups to become more directed into the water phase. This probably occurs as a result of the slightly distorted helix pairs interrupting the head group hydrogen bonding network more than the anti-parallel helix pair.

Therefore, the bilayer structural results indicate that the bilayer is stable and other than the altered acyl chain dynamic that occur due to helix pair inclusion, it represents a good model membrane-peptide environment

7.6. Discussion on Helix Structural Data

The helix structural data for the peptides again highlighted interesting results, with the three simulations indicating different behaviour.

The parallel helix pair, Hex_Ala2p, showed a high degree of dynamic behaviour further into the helix core residues than in the single helix simulations (except for the charged aspartate simulation). The helices appeared to reduce their lipid exposed surface area and began to come together in response to the destabilising effect of the aligned dipoles (Fig. 7.9a). The dihedral torsion and hydrogen bonding data indicated that the two spatially related C-termini had reduced ideal α_R geometry, whilst the N-terminus of helix 1 was mobile and fraying the N-terminus of helix 2 was totally disrodered and frayed.

The anti-parallel helix pair, Hex_Ala2a, lost a surprising (Fig. 7.9b) amount of helix for an arrangement that is thought to be favourable, but this reduced helicity was mainly at the N-terminal sections and could be deemed to be within the normal dynamic behaviour of a peptide helix pair in a polar environment. Changes in the structure of the C-termini of the two helices was observed, according to the dihedral torsion angles and the hydrogen bond results, but they indicate reduced mobility and fraying with respect to the N-termini.

The anti-parallel helix pair that had residues mutated to Trp/Tyr again showed dynamic behaviour, with the helices reducing their separation through cooperative changes in structure. Again the terminal regions were mobile and fraying but unlike the other two helix pair simulations the Hex_Trp_Tyr helices have weaknesses at specific sites away from the termini, which appear to allow the helices to achieve a reduced separation through a slight coiling effect. Interestingly, the helices appear to distort to accommodate the conformational requirements of the aromatic side chains (Fig. 7.9c), this will be investigated further in Chapter 8.

Again, there was little or no evidence of 3_{10} helical geometry being present in all the helix pair peptides, but the helices possess the non-α_R geometry observed in Chapter 6 for the single helix systems.

Thus, the bilayer and helix results indicate that these types of systems are stable equilibrated, and can produce detailed information on the interactions of a bilayer and peptide helix pairs. The peptides possess reduced α helicity from the single helix examples but their dynamic structures are stable. These systems will therefore allow the investigation and quantification of the interaction energies (Chapter 8) with the lipid bilayer and can be used to look at the driving force for bundle formation.

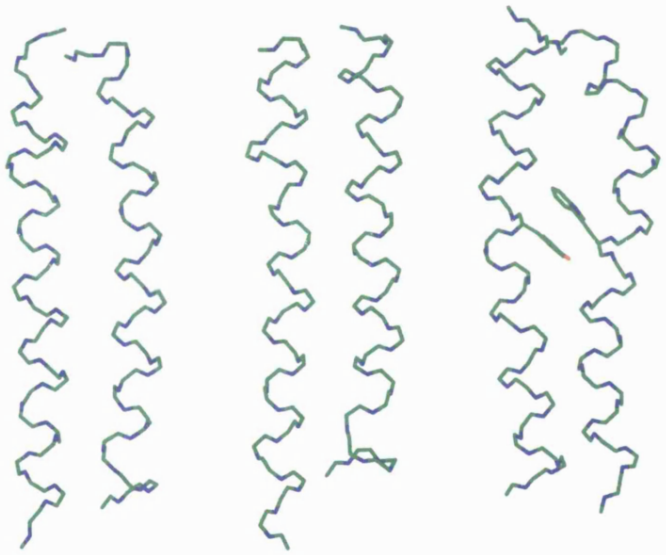


Fig. 7.9 - Helix Pair 100 ps Structures

References

1. L.Y. Shen, D. Bassolino, and T. Stouch, *Biophysical Journal*, **73**, 3-20 (1997).

8. Helix Partial Residue Non-bond Energies

8.1. Introduction

The helix stability results from Chapters 6 and 7 were considered and it was decided to take the helical section including residues 5-28 to calculate the average central helix alanine partial residue energy. Thus, the energies for residues 5-16 and 18-28 in the series of mutated polyalanine helix simulations and residues 5-28 in the pure polyalanine helix simulations were averaged to give the 5-28 average. The 5-28 average alanine value was calculated for each simulation, with the mutation site residue 17 partial residue energies being calculated separately. As well as the energies for the central residue of the helices the partial residue energy for the surface associated mutation sites at residues 4 and 29 was also calculated. These energetic results will allow the structural data described above to be analysed and the dynamic behaviour of the helix to be explained with respect to these energies that underpin the function of the peptide.

The average energies will be used to look at the interaction of the helices with the solvating lipid bilayer. How do the mutated amino acid side chains interact with the hydrophobic bilayer interior and the polar surface region of the bilayer? In order to facilitate the analysis of the partial residue energies, various atomic or residue level figures will be used to illustrate the basis for a result or to confirm a postulated mechanism of action.

The average energies will also be divided into intramolecular and intermolecular energies. The intramolecular energies will allow the intrinsic structural behaviour and stability of the helices to be analysed, the intermolecular energy will allow the analysis of the solvation enthalpy of the residues and the helices in the lipid bilayer. The solvation enthalpy of the mutated residues, in the different sections of the bilayer, is also analysed with respect to the change in structure and therefore, the intramolecular

energies.

The results for the partial residue energies will be presented in two sections, the first for the single helix incorporated system simulations and the second the helix pair system simulations. The results overall will then be discussed together in order to make conclusions about the energetics of the mutated residues.

8.2. Single Helix Simulation Results

8.2.1. Polyalanine Helix

The average partial residue energies for the Hex_Ala1 system are contained in Table 8.1 and the plot of the component averages across the simulation in Fig. 8.1.

Table 8.1

Hex_Ala1 - Helix Partial Residue Energies 5-28 Averages (kcal mol^{-1})			
	Total	V.D.W.	coulombic
Total	4.92	-0.73	5.66
Intramol.	9.11	2.72	6.39
Intermol.	-4.19	-3.45	-0.74

The partial residue total, intra and inter-molecular energies can be seen to equilibrate at different rates in Fig. 8.1. The helix stability results indicated that the helix relaxes into its dynamic structure after about 5-10 ps, i.e. after the warm up and template forcing phase of the simulation. From this point on the intra-molecular energies indicate that the helix is stable and dynamic. The intermolecular energies are equilibrated sooner, thus indicating that the lipid bilayer undergoes no major change in structure, therefore the interaction of the helix with the bilayer is stable.

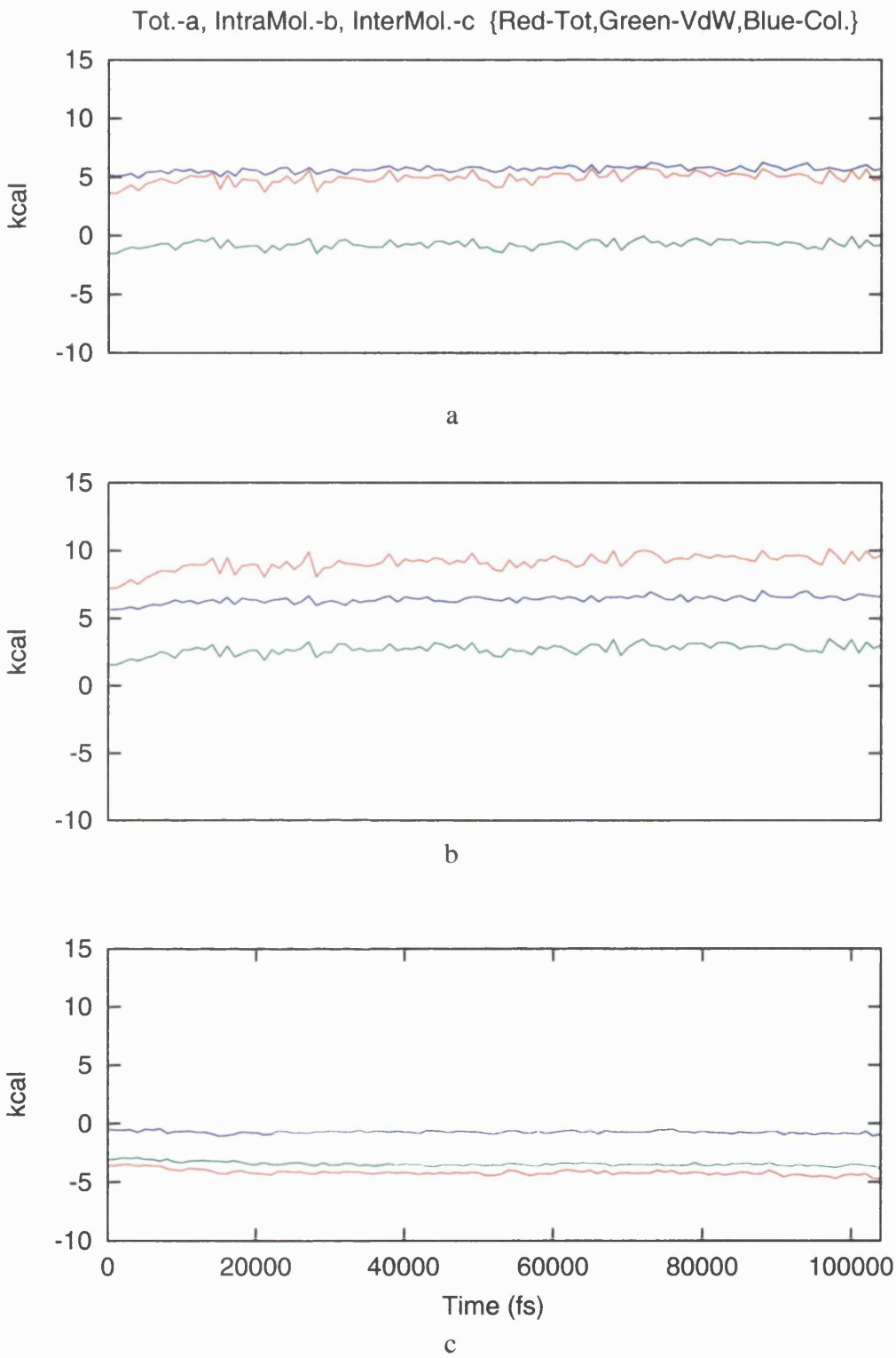


Fig. 8.1 - Hex_Ala1 Partial Residue Energies

As the peptide has already been proven to be in a good α_R helical structure its intermolecular partial residue energy represents the value for an alanine in an α_R helical peptide. The intermolecular average has a favourable solvation enthalpy in the bilayer with respect to the helix in vacuo.

The stabilising component in the intermolecular energy is the favourable Van der Waal energy indicating that the packing of the helix with the lipid bilayer is good and contributes more to the total intermolecular energy than the electrostatic energy. Thus the aliphatic nature of the C_β side chain allows favourable interactions with the hydrophobic core region of the bilayer.

8.2.2. Charged Aspartate Mutated Polyalanine Helix

The partial residue energies for the charged aspartate mutated polyalanine (Hex_Asp-) helix trajectory are contained in Table 8.2 and the plots of the values are illustrated across the simulation in Fig. 8.2.

The plots in Fig. 8.2 clearly illustrate the highly dynamic nature of the Hex_Asp- helix. The intramolecular energies are consistent with a mobile structure that undergoes change during the simulation and with the previously discussed helix structural stability results (Chapter 6).

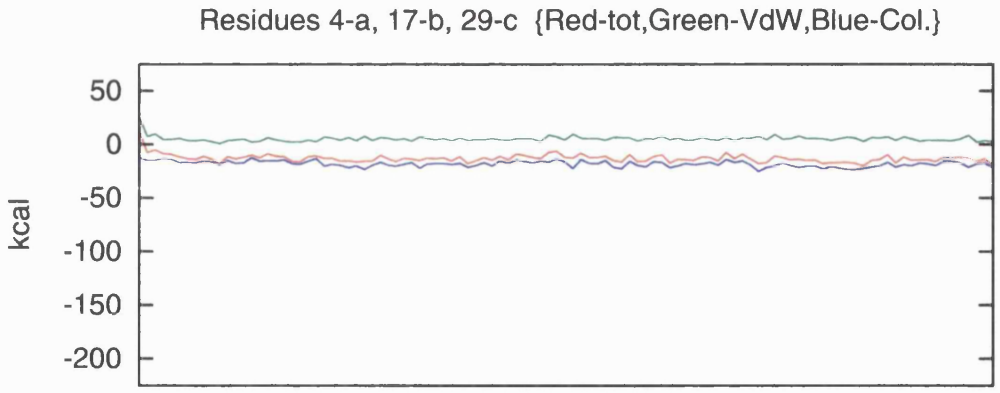
The average intramolecular residue energies for the aspartate residues illustrate the difference in their local geometries as discussed earlier in the thesis, i.e. the loss of helicity in both terminal sections of the helix. The loss of structure causes the Van der Waals energy to increase due to the loss of ideal geometry. The main loss of energy is from the electrostatic energy of the bottom mutation site, which loses about 11kcal and is a result of the loss of the backbone hydrogen bonding at the C-terminus and unfavourable interactions with the exposed terminal carbonyl groups. The structure of residues 4 and 17 are similar as their intramolecular energies are similar. Obviously the more polar environment of residue 29, resulting from the fraying and subsequent

increase in solvent interactions, is reflected in the change of the electrostatic energies, as it satisfies its hydrogen bonding requirement externally.

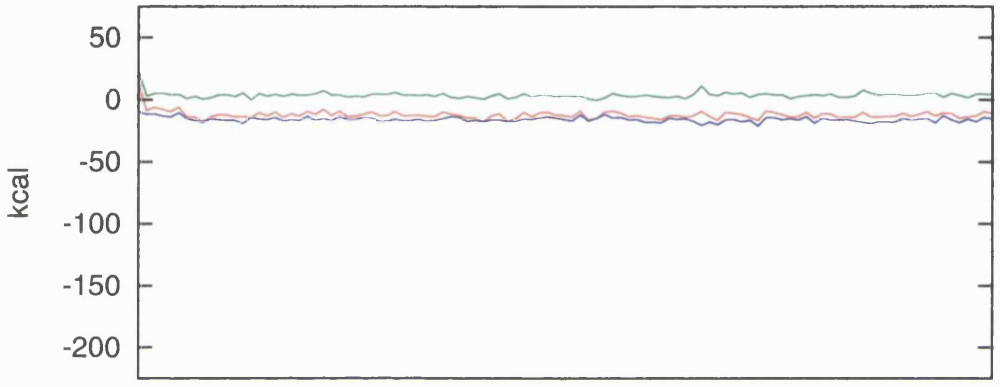
The intermolecular energy plots indicate that, whilst the intramolecular energy is equilibrated early in the simulation, the interaction of the terminal residues with the lipid bilayer is not stable. The central aspartate residue is clearly reasonably stable throughout the simulation but the fact that the two terminal intermolecular energies are unstable agrees with the observed mobility of the termini. The intermolecular value for the upper mutation site, residue 4, remains closer to the original value, with only a dip between 40 and 60 ps. This change is a result of the temporary loss of ideal structure and the elongation of the hydrogen bond for the residue, as observed in hydrogen bond results (see Supplementary Work, Chapter II). The graph for the bottom mutation site residue 29, indicates that the gain in interaction energy occurs at 20 ps and is a result of the conformational changes in the helix at this point in the simulation. From this point in the simulation the loss of the terminal structure removes the block to water molecules approaching the Asp- side chain and thus improves the electrostatic environment of the carboxylate functional group.

Table 8.2

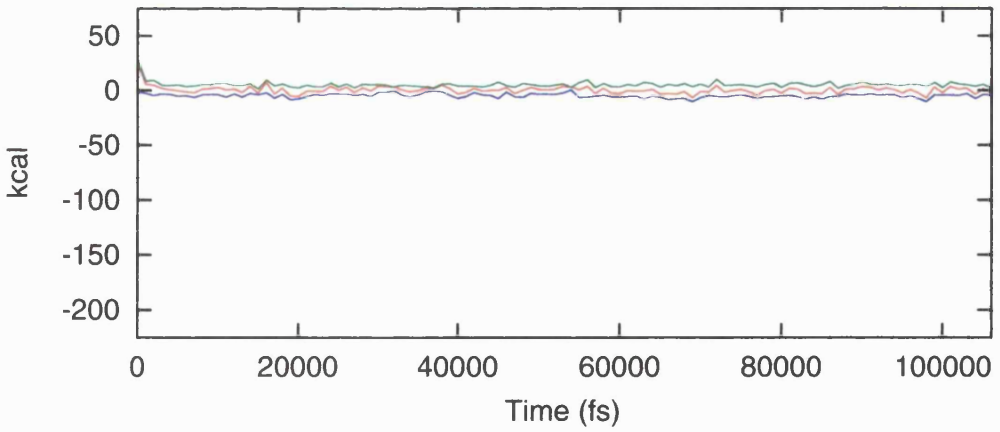
Hex_Asp-, Helix Residue Partial Energies (kcal mol^{-1})									
	Residue 4			Residue 17			Residue 29		
	total	V.D.W.	coul	total	V.D.W.	coul	total	V.D.W.	coul
Total	-84.18	2.86	-87.05	-11.05	-1.35	-9.70	-89.74	2.57	-92.31
Intramol.	-12.98	5.21	-18.19	-12.04	3.83	-15.86	0.39	5.15	-4.76
Intermol.	-71.20	-2.35	-68.86	0.99	-5.18	6.16	-90.13	-2.58	-87.55
				Alanine Residue 5-28					
Total				5.11	-0.58	5.69			
Intramol.				9.40	2.81	6.58			
Intermol.				-4.28	-3.38	-0.89			



a

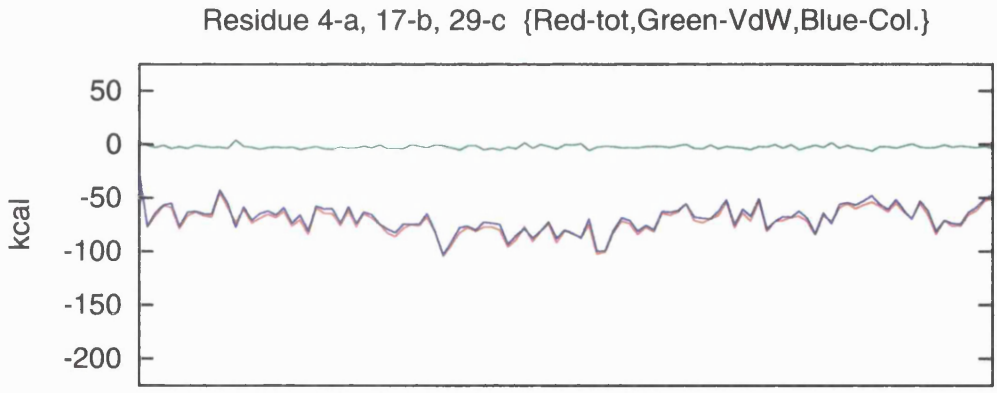


b

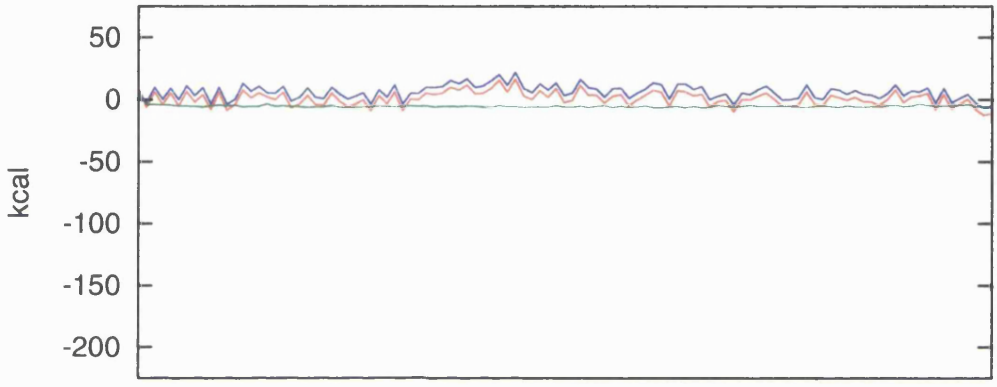


c

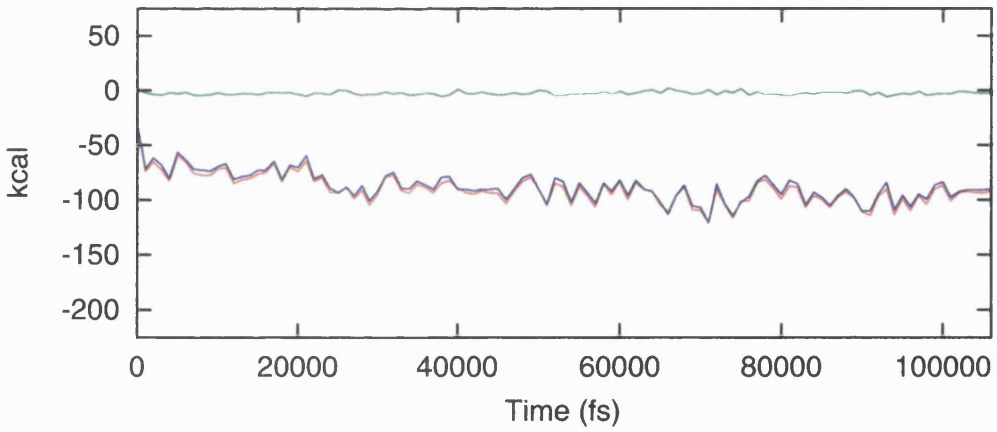
Fig. 8.2 - Hex_Asp- Intramolecular Partial Residue Energies



a



b



c

Fig. 8.2 - Hex_Asp- Intermolecular Partial Residue Energies

Overall, the intermolecular data indicates that the top aspartate residue remains in the same general environment and structure whilst the lower aspartate undergoes a change in its environment as a consequence of conformational changes in the helix. This explains the -71.2 to -90.13 differential between the intermolecular energy terminal residue averages. The intermolecular energy for the middle mutation site, residue 17, indicates that this helix undergoes more non-ideal motion than most of the helices, with its profile fluctuating throughout the simulation and gaining a small amount of energy in the middle section. No physical reason for this can be seen other than it resulting from the mobility of the helix core.

The intermolecular energy of residue 4 represents the solvation energy of a helical charged aspartate in the interface region of a bilayer, whereas the intermolecular energy of the residue 29 represents the solvation energy for a less ideal helical charged aspartate residue in a bilayer environment. Thus the solvation energy for a helical aspartate residue in a lipid bilayer environment indicates that the inclusion of this type of residue in a membrane peptide is highly unfavourable with respect to inclusion in the interface region, (-71.20 or -90.13 kcal mol^{-1}) and thus, could be associated with peptide or protein helical anchoring at the interface of the bilayer.

The alanine residue 5-28 average energies indicate that the alanine residues in the core of the helix have a different structure to Hex_Ala1 and thus confirms the previous helix stability results (Chapter 6). The lower stability of the Hex_Asp- helix core is reflected in the slightly higher intramolecular energy of the alanine average, which results from a slightly higher VDW energy indicating a poorer packing of the atoms and from an unfavourable electrostatic energy change indicating a reduced hydrogen bonding stabilisation contribution. Thus, the charge of the side chain of residue 17 causes a destabilisation of the helix core alanine residues by 0.29 kcal mol^{-1}.

The average intermolecular energy of the 5-28 alanines indicates that the interactions of the alanine core residues are slightly more favourable than for the Hex_Ala1

system simulation, but the change is small. The VDW energy suggests the mobile nature of the Hex_Asp- helix has caused a reduction in the packing stability and the increased electrostatic stabilisation could be due to the changed electrostatic interactions in the upper regions of the helix.

The charged aspartate side chain of residue 17 begins the simulation in an extended conformation (Fig. 8.3a) with the charged side chain end close to its backbone carboxyl oxygen, thus destabilises the backbone. During the simulation the side chain minimises this interaction by looping around to closer to its amide hydrogen. This is an intermittent event but does explain the destabilisation of the helix geometry, as such a close approach of a charge will effect the hydrogen bonding of the structure over a number of residues. Also, the change in local geometry will place strain on the adjacent residues.

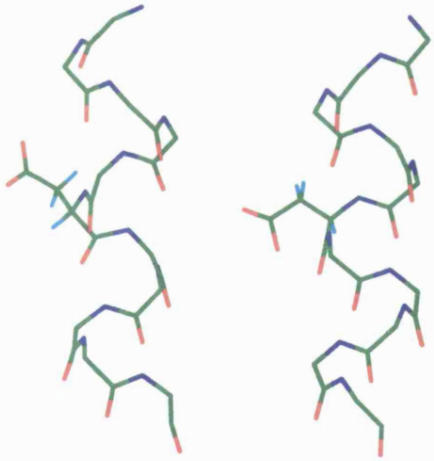


Fig. 8.3a Hex_Asp- Residue 17 Side Chain

The destabilisation of the C-terminus and the increased intermolecular stabilisation of the aspartate residue 29 can be explained by a bridging water between residues 25 and 29, illustrated in Fig. 8.3b (the figure includes two views of the same side chain water section of the system).

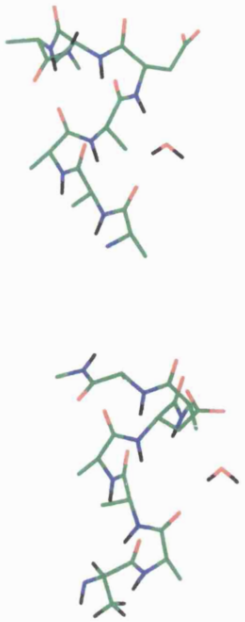


Fig. 8.3b Hex_Asp- Residue 29 Side Chain Bridging Water

8.2.3. Neutral Aspartic Acid Mutated Polyalanine Helix

The partial residue energies for the neutral aspartic acid mutated polyalanine helix (Hex_Asp) simulation are contained in Table 8.3 and the plots of the values are illustrated across the simulation in Fig. 8.4.

The plots of the total energies clearly indicate that the Hex_Asp helix is stable and undergoes no major changes during the simulation, which is in agreement with the structural results.

The intramolecular energies indicate that there is a difference in structure, between the three residues, with the two termini exhibiting different properties. This again is in accordance with the structural results. The intramolecular electrostatic energies of residue 17 is more negative and the cause of this increased electrostatic stability is the hydrogen bonding of the aspartic acid side chain with its backbone carbonyl oxygen, see Fig. 8.5.

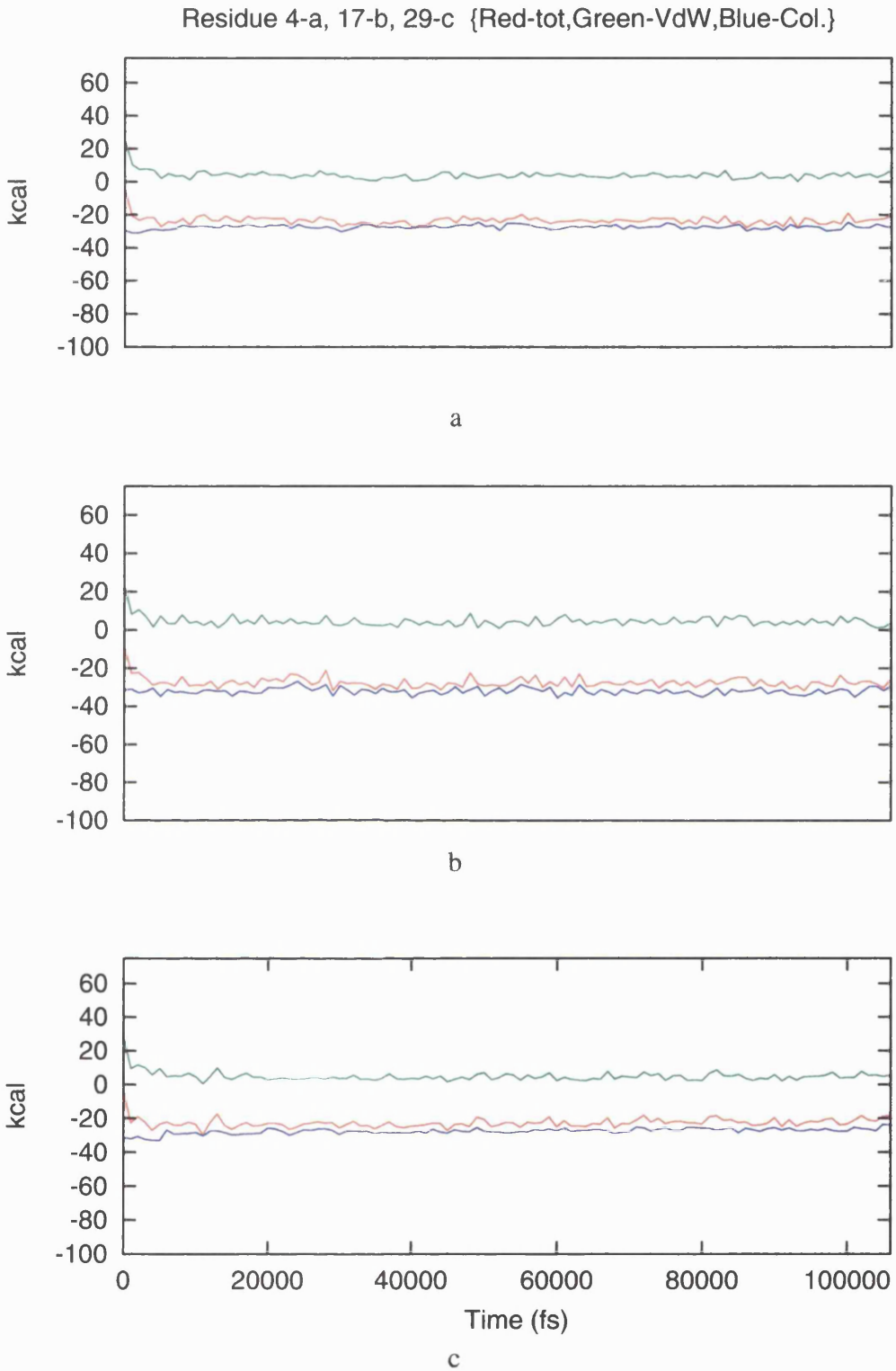


Fig. 8.4 - Hex_Asp Intramolecular Partial Residue Energies

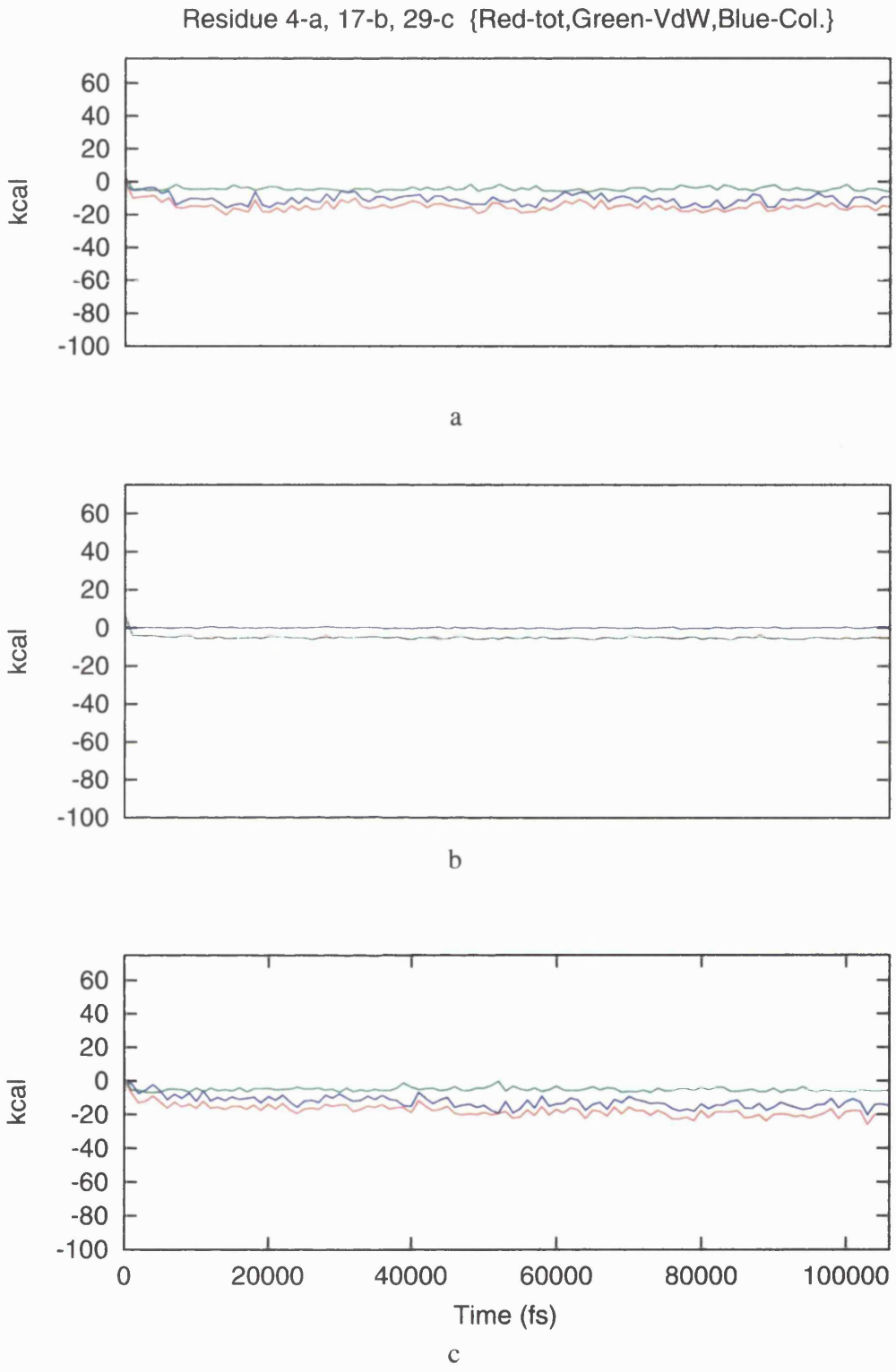


Fig. 8.4 - Hex_Asp Intermolecular Partial Residue Energies

The intramolecular electrostatic energy indicates that residues 4 and 29 have identical hydrogen bonding, but the stability data indicates that the N-terminus is frayed but it does not propagated down to the aspartate residue. The aspartic acid group does not loop around to the backbone in the terminal region therefore has no effect on the terminal residue stability indicating that the N-terminal fraying is a result of a intermolecular event and not an intrinsic effect of the mutated residue 4.

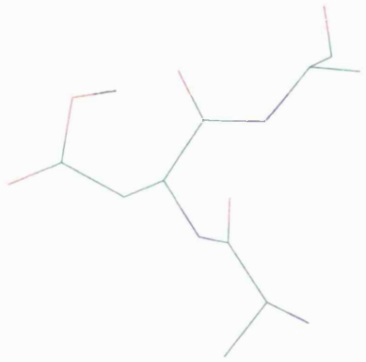


Fig. 8.5 Hex_Asp Residue 17 Side Chain

The intramolecular VDW energy of the three mutation sites reflects the changes in structure, with the lower value of the residue 4 average indicating a more open structure at the N-terminus and the slightly higher value for residue 29 indicates the slightly more compressed nature of the helix geometry.

Table 8.3

Hex_Asp, Helix Partial Residue Energies (kcal mol^{-1})									
	Residue 4			Residue 17			Residue 29		
	total	V.D.W.	coul	total	V.D.W.	coul	total	V.D.W.	coul
Total	-38.40	-0.29	-38.12	-32.44	-0.57	-31.87	-40.21	-0.22	-40.00
Intramol.	-23.40	4.04	-27.45	-27.44	4.52	-31.96	-22.64	4.89	-27.53
Intermol.	-15.00	-4.32	-10.67	-5.00	-5.09	0.09	-17.58	-5.11	-12.47
				Alanine Residue 5-28					
Total				4.90	-0.70	5.61			
Intramol.				8.91	2.70	6.22			
Intermol.				-4.01	-3.40	-0.61			

The intermolecular energy averages in Table 8.3 indicate a slightly altered local environment structure between the termini and the core residues and a more negative electrostatic energy in the residue 29 average when compared to residue 4. The plot for the intermolecular energy for residue 4 indicates that there is a two phase profile to the plot, a slightly higher energy is observed in the template forced period of the simulation and then a generally lower and fluctuating period until the end of the simulation. The energy gets more positive towards the 30 ps mark in the simulation which indicates that the COOH group is loosing hydrogen bonding as there is no compensating change in the intramolecular energies. The 30 ps point in the simulation is when the structure of the N-terminus is lost through a transition to a more extended conformation.

The helix structural data categorised the core of the helix as being in a good α_R geometry and thus the intermolecular energy of residue 17, -28.44 kcal mol^{-1}, is the solvation energy of a aspartic acid amino acid in a DMPC bilayer. The intermolecular energy of residue 29,-15.00 kcal mol^{-1}, represents the energy of an α_R helical aspartic acid residue in the interface region of the lipid bilayer, whereas the intermolecular energy for residue 4,-17.58 kcal mol^{-1}, represents the solvation energy of an unfolded-ing α_R helical aspartic acid, in the interface region.

The solvation energy for the helical aspartic acid is only slightly less favoured (3:1) in the core region than at the interface region of the bilayer, when it is compared with the Hex_Asp- example, which is very highly favoured at the interface of the bilayer. Thus, the aspartic acid residue can be argued to favour both positions in the lipid bilayer but with a preference for the interface region as would be expected for a hydrophilic body.

The 5-28 alanine average energies indicate that the helix alanine core is slightly more stable than the core in the pure alanine simulation by 0.2 kcal mol^{-1}. The VDW energy, as it is slightly lower than for the Hex_Ala1 system, indicates that the packing

of the structure is slightly improved in the Hex_Asp simulation due to the more ideal helical structure. The intermolecular energy of the alanine core is slightly higher than the pure helix by 0.18 kcal mol^{-1}. This indicates that the more ideal helix has a poorer VDW interaction of the C_β side chain with the lipid bilayer.

8.2.4. Arginine Mutated Polyalanine Helix

The partial residue energies for the Arginine Mutated Polyalanine Helix (Hex_Arg) simulation are contained in Table 8.4 and the plots of the values are illustrated in Fig. 8.6.

The plots of the intramolecular average energies in Fig. 8.5 are equilibrated, which is in agreement with the helix structural data. The intramolecular energetics of the central mutation site exhibit a shallow dip in the electrostatic energy at about 40 ps. This indicates that a change occurs, during this period of the simulation, in the structure of this residue. No changes in energy are observed at this point in the simulation for residues 4 and 29.

The plot of the intermolecular energies of the mutated residues indicates that the interactions of the two terminal mutated residues are dynamic. The major feature of these plots is that there is a significant change in the intermolecular interactions of residue 29 at about 30 ps and at 60 ps. The changes of energy reflect the hydration of the two termini and the reaction of the arginine residues. The side chain of residue 4 has a stable interaction with the surrounding waters and remains near to its original structure. The side chain of residue 29 starts the simulation directed away from the water layer and also, experiences a reduced level of hydration. It obviously attempts to optimise its interactions with the reduced hydration it experiences (Fig. 8.9), but it is unable to totally achieve this stabilisation. The plots confirm this as it has a final energy a little below that of the side chain of residue 4.

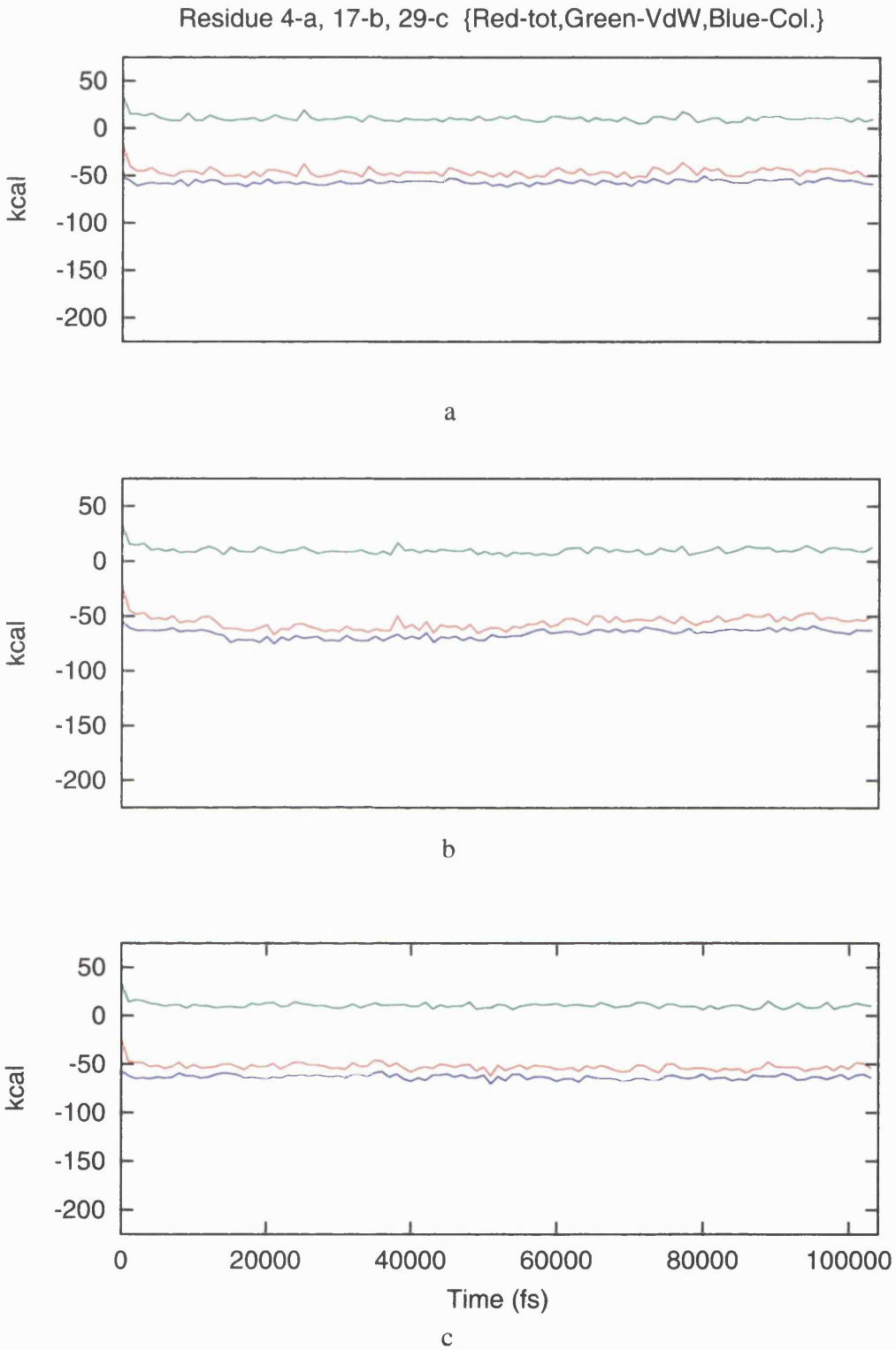


Fig. 8.6 - Hex_Arg Intramolecular Partial Residue Energies

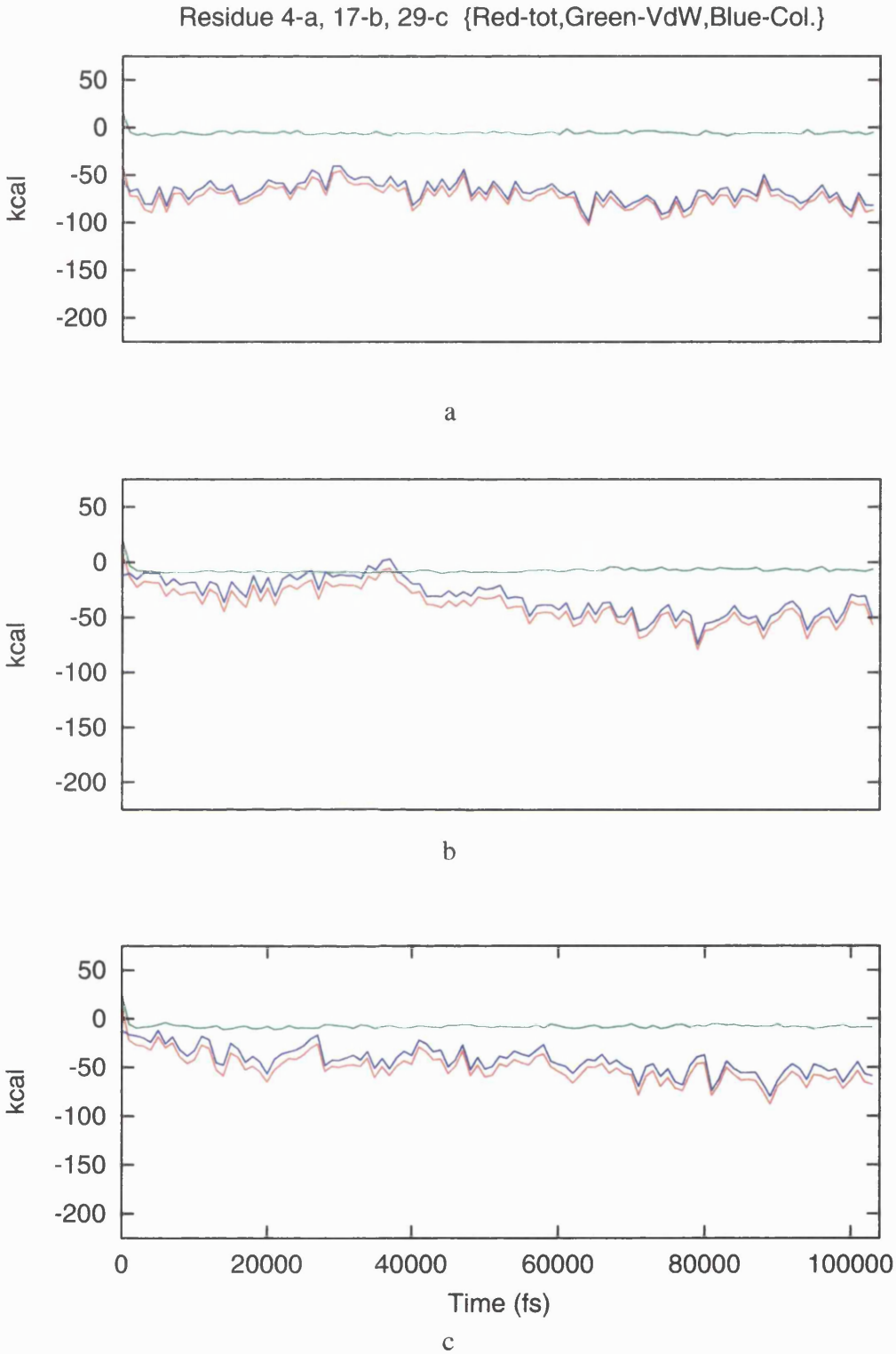


Fig. 8.6 - Hex_Arg Intermolecular Partial Residue Energies

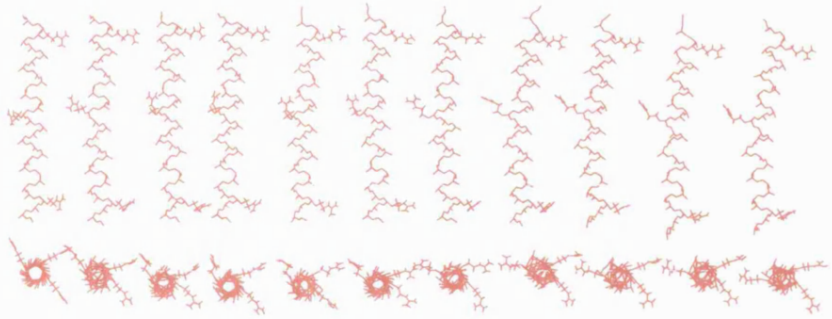
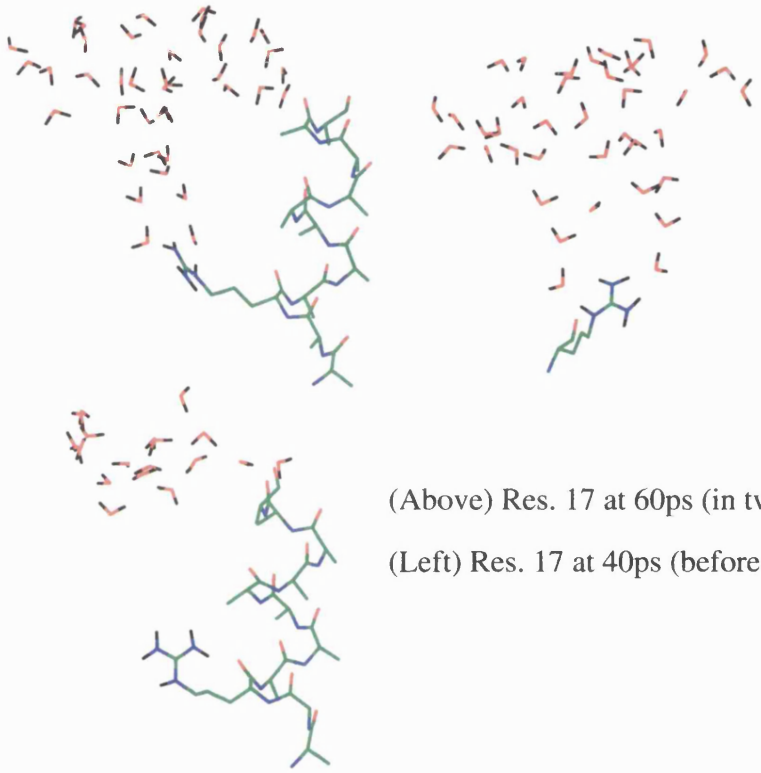


Fig. 8.7 Hex_Arg Helix - Long Axis and Plan (0 - 100 ps every 10 ps)

The intermolecular energies of the central mutation site (Fig.8.6b) suggest as did the intramolecular data that a major conformational change occurs, resulting in the side chain experiencing an altered environment.

The time elapsed coordinate diagram in Fig. 8.7 indicates that the side chain of residue 17 changes its structure as the simulation progresses with the side chain changing from being directing into the lipid core to being directed towards the surface of the bilayer. Fig. 8.8 includes a representation of residue 17 together with water molecules that are within 10 Å of the COOH entity both before the observed transition at 40 ps and after 60 ps. Clearly in the initial section of the simulation the side chain is looped around and hydrogen bonding to the backbone, which is the reason for the increased negative intramolecular energy during the period where this hydrogen bond is present, 20-60 ps. The second arrangement is very interesting and indicates that the arginine residue flips its structure to utilise its length and maximise its interactions through hydrogen bonding with the bilayer surface waters. The waters begin to 'feel' the attraction of the arginine at about 40 ps and then the waters are drawn gradually through the bilayer until they become associated directly with the arginine. This

interesting observation indicates that the amino acids in the bilayer core can stabilise their polar interactions, if they have a long side chain, by 'reaching up' to the surface waters. This new arrangement is stable throughout the simulation. The average intermolecular energy of residue 17 is -38.72 kcal mol^{-1} but the values for the two phases are -24.01 kcal mol^{-1} for the first period and -51.11 kcal mol^{-1} for the second period of the simulation.



(Above) Res. 17 at 60ps (in two orientations)

(Left) Res. 17 at 40ps (before transition)

Fig. 8.8 - Hex_Arg Residue 17 and a 10 Å Radius Shell of Waters

Thus, the intermolecular energy of residue 17 represents the solvation energy of the arginine α_R helical residue in a bilayer environment in two different types of side chain structure and interaction, both conformations being negative. The intermolecular energy of residues 4 and 29 represent the solvation of the helical arginine residue in the bilayer surface region but the two observed values are a result of the differential hydration of the side chains (Fig. 8.9), which is a result of the structure of the bilayer, an artefact of the building and the way the SOAK programme soaked the system. The

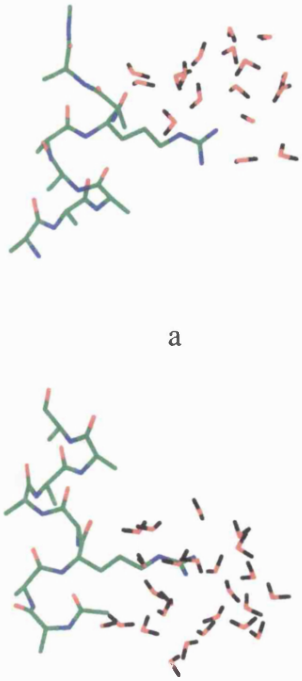
frayed C-terminus is not a results of the arginine side chain behaviour and occurs at the C-terminus despite the N-terminus being more hydrated and proceeds via competitive water attack.

The centrally located arginine residue when in the second conformation, i.e. with the side chain directed into the surface waters, is seen to be as stable as an arginine residue at the surface with partial hydration, therefore the polar side chain of the arginine can achieve a solvation arrangement that will make bilayer core bound arginines as likely to occur as a bilayer surface bound arginines. This is against accepted wisdom that such a charged entity is likely to occur in the bilayer interior. Also, the negative value of the intermolecular energy before the transition in side chain structure indicates that the charged arginine side chain has a good interaction with the lipid bilayer interior.

Table 8.4

Hex_Arg, Helix Partial Residue Energies (kcal mol^{-1})									
	Residue 4			Residue 17			Residue 29		
	total	V.D.W.	coul	total	V.D.W.	coul	total	V.D.W.	coul
Total	-119.20	4.78	-123.98	-94.45	2.81	-97.25	-103.25	2.83	-106.08
Intramol.	-46.28	10.42	-56.70	-55.73	10.13	-65.86	-52.71	10.62	-63.33
Intermol.	-72.92	-5.64	-67.28	-38.72	-7.33	-31.40	-50.54	-7.80	-42.74
				Alanine Residue 5-28					
Total				4.82	-0.68	5.50			
Intramol.				8.70	2.63	6.06			
Intermol.				-3.88	-3.31	-0.57			

The 5-28 alanine average energies in Table 8.4 indicate that the core helix structure is more stable than the Hex_Ala1 simulation by 0.21 kcal mol^{-1}. The more ideal nature of the helix geometry, as with the Hex_Asp simulation, results in a better packing of the helix atoms and therefore a slightly lower VDW energy. The main stabilisation comes from a lower electrostatic energy that results from the hydrogen bonding of the arginine side chain back with the backbone during the first part of the simulation.



a

b

Fig. 8.9 Hex_Arg Helix Hydration Around Residues 4(b) and 29(a)

(bilayer interior is located in the centre of the figure)

The intermolecular 5-28 alanine average indicates that the helix core has less favourable interactions with the bilayer than the Hex_Ala1 helix core. The favourable energetics of the helix must therefore come from intramolecular interactions.

8.2.5. Phenylalanine Mutated Polyalanine Helix

The partial residue energies for the phenylalanine mutated polyalanine helix (Hex_Phe) simulation are contained in Table 8.5 and the plots of the values are illustrated across the simulation in Fig. 8.10.

The plots for the intramolecular residue energies all indicate that the three phenylalanine residues are in stable conformations. Although the plots are equilibrated they do exhibit the fluctuations which are also observed in the helix stability data and are attributable to the dynamic motion of the system.

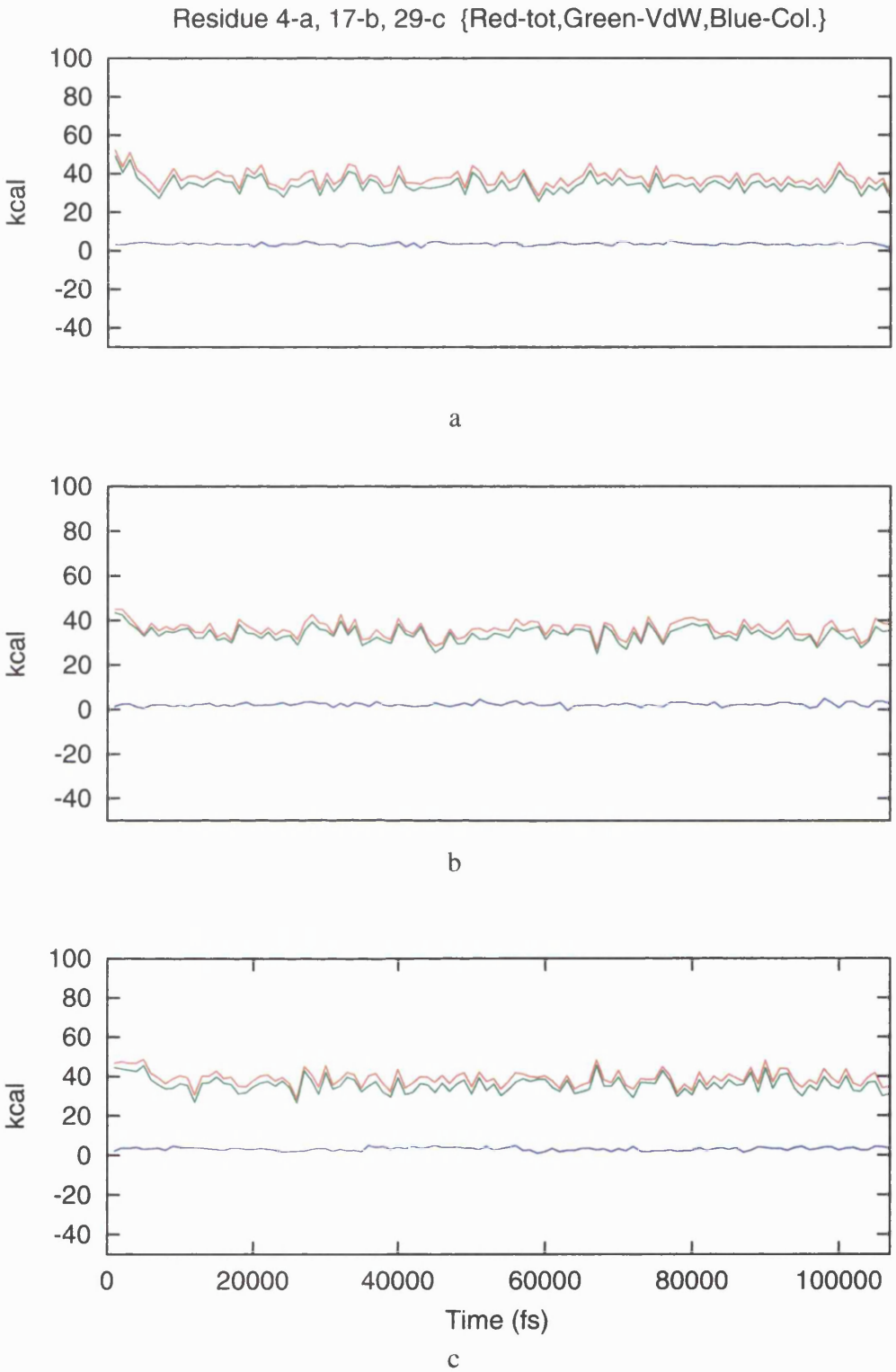


Fig. 8.10 - Hex_Phe Intramolecular Partial Residue Energies

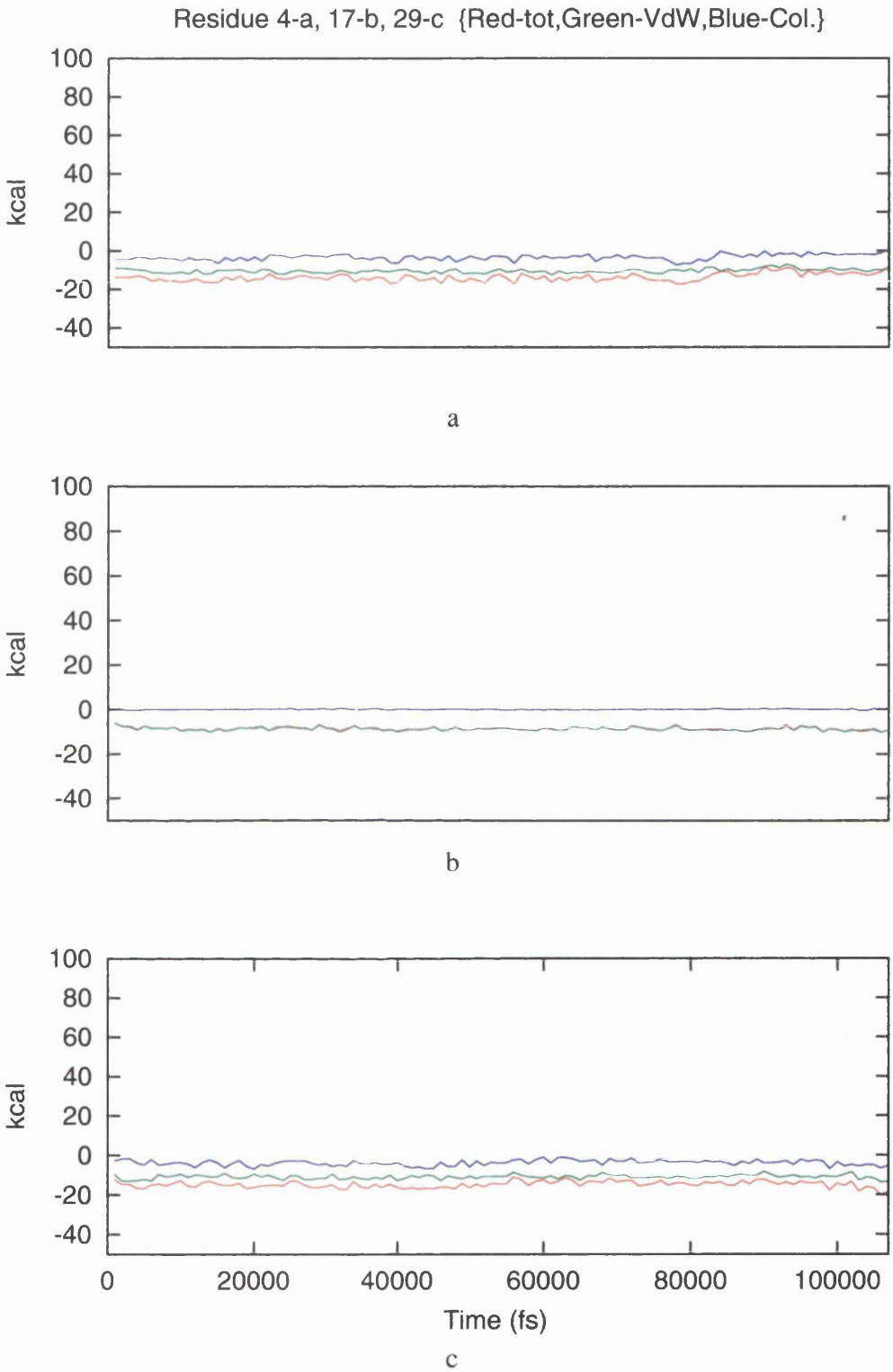


Fig. 8.10 - Hex_Phe Intermolecular Partial Residue Energies

The Chi torsion angles (Table 8.6) of the phenyl alanine residue have a slight variation but the χ_1 values for the three helix positions are relatively stable around a value of ≈170°

Table 8.5

Hex_Phe, Helix Partial Residue Energies (kcal mol^{-1})									
	Residue 4			Residue 17			Residue 29		
	total	V.D.W.	coul	total	V.D.W.	coul	total	V.D.W.	coul
Total	23.95	23.90	0.05	27.55	25.13	2.41	24.15	24.81	-0.65
Intramol.	37.78	34.35	3.43	36.16	33.85	2.34	38.97	35.78	3.19
Intermol	-13.82	-10.45	-3.38	-8.61	-8.72	0.11	-14.82	-10.97	-3.85
				Alanine Residue 5-28					
Total				4.89	-0.72	5.61			
Intramol.				8.88	2.60	6.28			
Intermol.				-3.99	-3.32	-0.67			

The intermolecular partial residue energies for the Hex_Phe system indicate that the three mutation sites possess stable interactions with the lipid bilayer, with the two terminal residues exhibiting similar values. These values are interesting as the phenyl ring is commonly designated as a hydrophobic entity and is thought to populate the lipid bound region of a protein and but is also thought to be an 'anchor' for helices to the membrane interface region. These energetics results are in general agreement with this position. Clearly a helical phenylalanine amino acid has a favourable interaction with both position in the lipid bilayer.

Looking at the component parts of the partial energies it can be seen that the phenyl ring is closer and better packed in the interface region and has better electrostatic interactions by about 4 kcal in the interface region than the core region, where the electrostatic energy is essentially zero. This is as would be expected for the electron dense ring interacting with the polar head group layer of the lipid bilayer. Therefore, the Phe aide chain prefers the interface region to the interior of the bilayer.

Therefore, the solvation energy for the phenylalanine amino acid in a α_R helical

geometry in the acyl region of a lipid bilayer environment is represented by the intermolecular partial residue energy of residue 17, -8.61 kcal mol^{-1} and the solvation energy in the interface region is represented by the intermolecular partial residue energy of residues 4 and 29, -13.82 and -14.82 kcal mol^{-1}.

Table 8.6

Hex_Phe, Side Chain $\chi_{1,2}$ Angles						
	Residue 4		Residue 17		Residue 29	
Time	χ_1	χ_2	χ_1	χ_2	χ_1	χ_2
0	-171.5	-173.3	-168.0	-164.8	-177.0	-162.9
10000	-177.0	-168.8	-179.9	-94.4	172.5	-122.0
20000	-158.7	-150.0	-178.5	-85.9	178.2	-163.8
30000	175.8	-143.3	-168.0	-100.6	167.2	-168.9
40000	-174.7	-153.7	179.9	-87.0	-149.1	-173.0
50000	-170.8	-150.8	-169.0	-119.3	-170.4	-149.0
60000	171.9	-124.2	-168.8	-118.4	166.1	-99.8
70000	178.1	-148.1	-172.0	-86.9	174.8	-138.2
80000	-177.9	-121.3	158.2	-102.9	179.6	-125.7
90000	-150.4	-118.8	-172.0	-124.2	-165.2	-111.9
100000	-174.2	-137.5	-165.1	-114.0	-178.0	-107.4

The 5-28 partial residue energy for the core alanine residues indicates that the structure of these residues is good α_R helical geometry, which is in agreement with the helix stability results in indicating the Hex_Phe helix being stable. Both the VdW and the electrostatic energies are slightly lower in energy indicating a better geometry and a more ideal hydrogen bonding network as a result of the better packed structure. Again, as observed for other helices which exhibit a more stable structure than the Hex_Ala1 simulation, the more ideal helical structure of the Hex_Phe system means the intermolecular energy is slightly below the Hex_Ala1 value. Thus, the interaction of the helix with the lipid bilayer is reduced as the helix is more compact and satisfies its stability requirement internally.

8.2.6. Threonine Mutated Polyalanine Helix

The partial residue energies for the threonine mutated polyalanine helix (Hex_Thr) simulation are contained in Table 8.7 and the plots of the values are illustrated across the simulation in Fig. 8.11.

The plots of the partial residue energies indicate that the three Thr residues are in an essentially stable structure and experience a constant environment. The intramolecular energy plots indicate that, following the initial template forced 5 ps of the simulation, the energies settle to an equilibrium and exhibit an oscillatory nature around their equilibrium value. The intermolecular energy plots also indicate that after the initial period the energies are stable, oscillating around an equilibrium value but there is evidence that the threonine at residue 4 undergoes a transition in its interactions. This change is stepped and acts to give residue 4 a similar equilibration energy to that of residue 29. The intermolecular partial residue energy for residue 17 is very stable, with an almost non-existent electrostatic energy, as would be expected for an entity in an apolar environment.

The average residue partial energies in Table 8.7 indicate that the threonine amino acid has an affinity for the three different helix positions in the bilayer. The intramolecular energy of residue 4 is slightly above that of residue 29, a result of a higher coulombic energy which is caused by an unfavourable interaction of the backbone CO' and the side chain OG atom. These two atoms are closer to each other in the structure of residue 4 than residue 29 and hence the increased electrostatic energy is a result of the repulsion of these two negative, partially charged atoms. The lower intramolecular energy for residue 17 is due to the better helical geometry observed in this region of the helix, with its more stable core residues and therefore better hydrogen bonding gives a lower electrostatic energy.

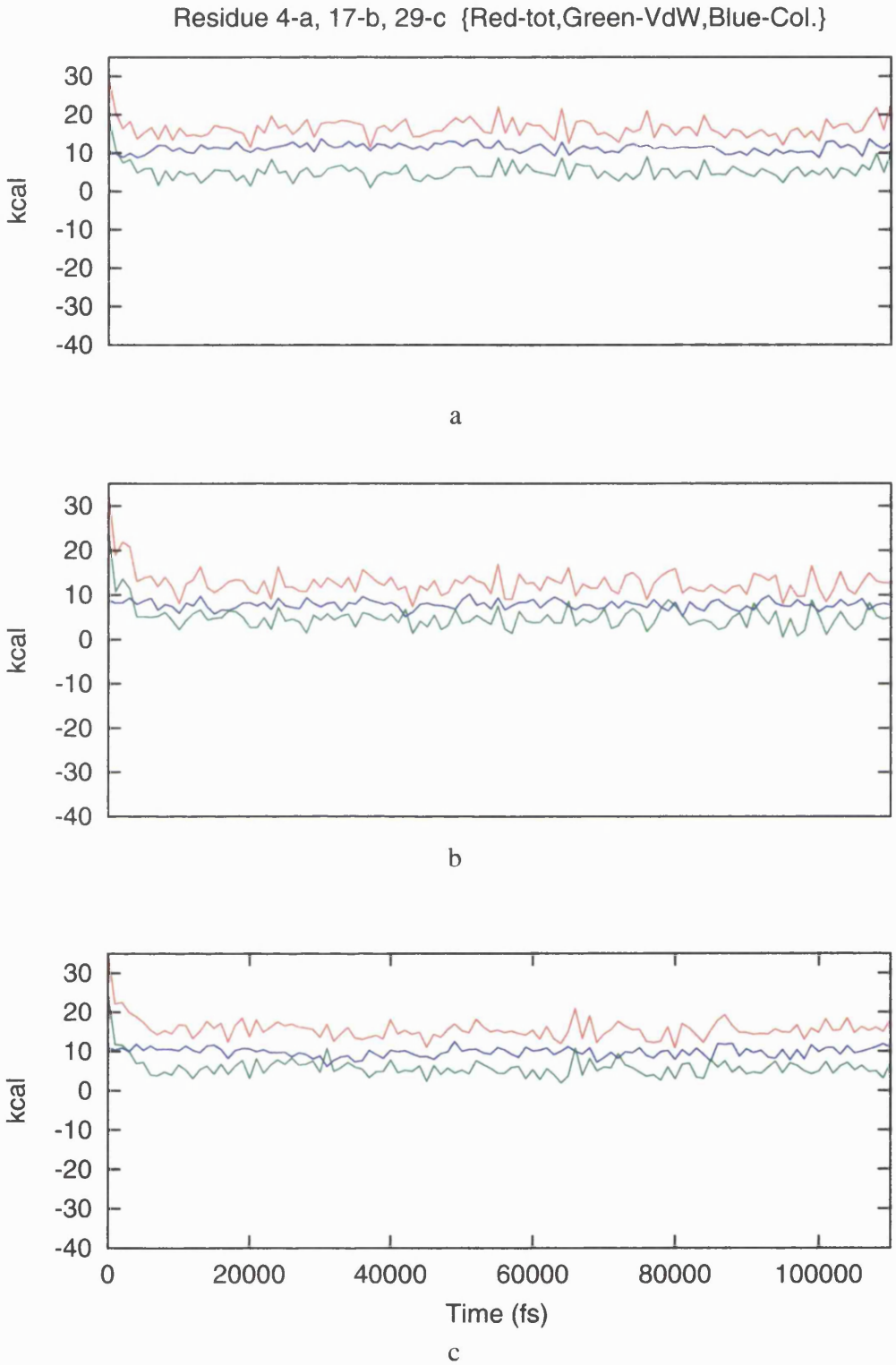


Fig. 8.11 - Hex_Thr Intramolecular Partial Residue Energies

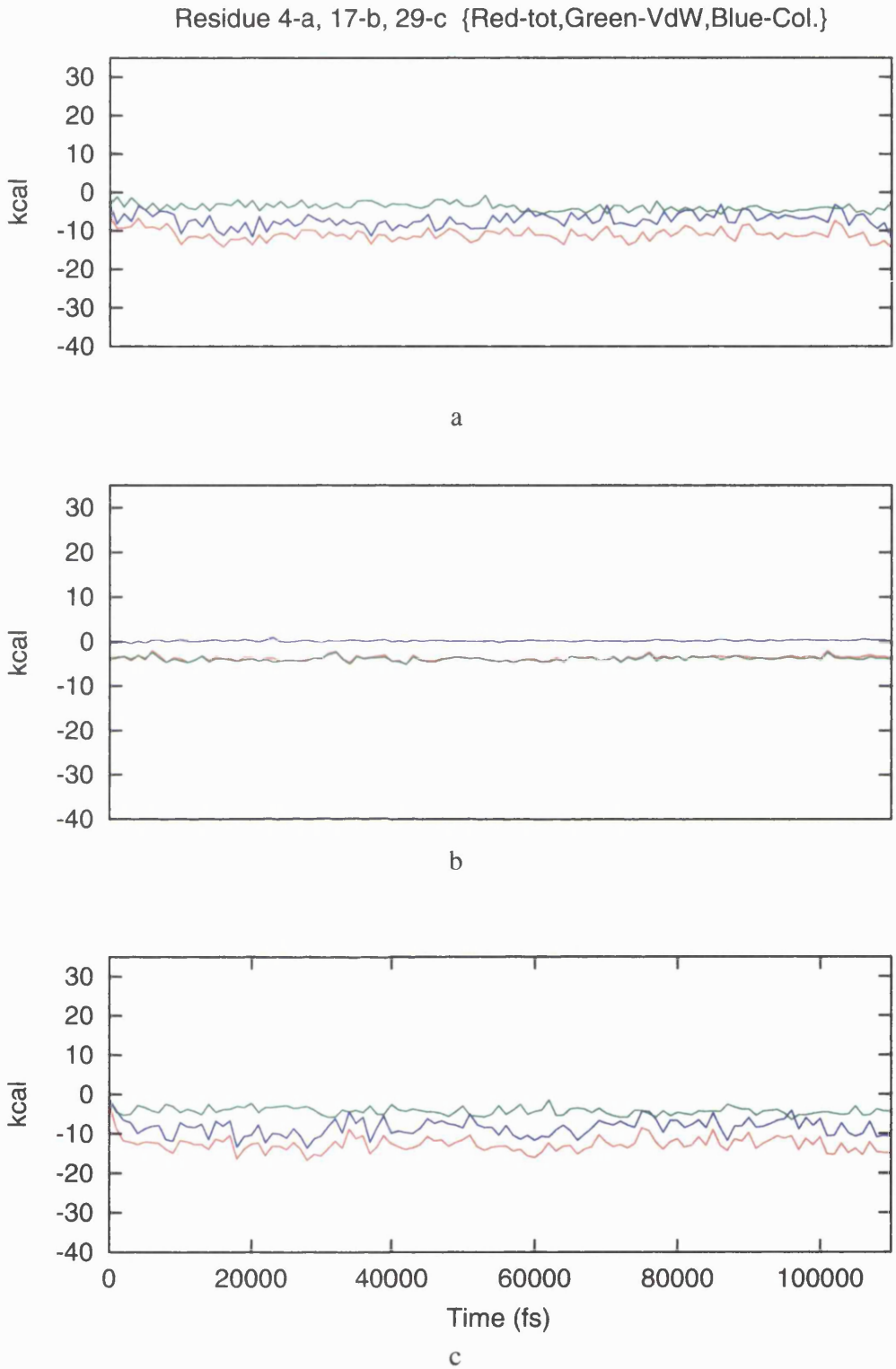


Fig. 8.11 - Hex_Thr Intermolecular Partial Residue Energies

The intermolecular energy averages indicate that the threonine amino acid is more stable in the interface region of the bilayer than the core region. The higher energy for the centrally located threonine is due to the non-polar nature of the bilayer interior, which is unable to satisfy the electrostatic requirements of the side chain polar atoms whereas the more polar surface of the bilayer can satisfy the side chain hydrogen bonding requirement. The threonine residue has a reasonably negative interaction energy in the interior of the bilayer, even though there is no hydrogen bonding possibility, because of the good packing of the small side chain with the lipid environment as illustrated by the favourable VDW energy. The difference in the terminal averages is due to the early phase of the simulation where residue 4 is higher in energy until the transition at 4 ps, whereas residue 29 more efficiently and gradually achieved its conformation. This result is a consequence of the higher C-terminal fluidity observed in the helix structural results, giving a more ideal packing of the side chain, i.e. lower VdW energy and also a lower electrostatic energy as the fraying allows these interactions to be better satisfied. Interestingly the intramolecular energy also is affected by the fraying of the C-terminus because although the VdW energy increases as the better packing of the helix is lost the loss of helical structure relieves the OG-CO interaction, thereby giving a lower electrostatic energy.

Table 8.7

Hex_Thr, Helix Partial Residue Energies (kcal mol^{-1})									
	Residue 4			Residue 17			Residue 29		
	total	V.D.W.	coul	total	V.D.W.	coul	total	V.D.W.	coul
Total	5.47	1.42	4.05	9.06	1.06	7.99	2.94	1.59	1.35
Intramol.	16.47	5.21	11.26	12.74	4.94	7.80	15.68	5.85	9.83
Intermol.	-11.00	3.79	-7.21	-3.69	-3.88	0.19	-12.75	-4.27	-8.48
				Alanine Residue 5-28					
Total				4.90	-0.72	5.62			
Intramol.				8.94	2.67	6.27			
Intermol.				-4.04	-3.39	-0.65			

Thus, the intermolecular partial residue energy of residue 4 represents the solvation energy for a threonine α_R helical N -terminal located residue in a DMPC bilayer surface, -11 kcal mol^{-1} and the average for residue 29 presents the solvation energy for a threonine partial frayed, α_R helical C -terminal located residue, -12.74 kcal mol^{-1}. The solvation energy for an α_R helical residue in the hydrophobic core of a DMPC bilayer is represented by the intermolecular energy, -3.69 kcal mol^{-1}.

The 5-28 average energies for the alanine helix core are contained in Table 8.7 and indicate that the structure of the residues is stable and more helical than the pure alanine Hex_Ala1 simulation (for the same reason formed earlier). This again indicates the inverse relationship between helix core stability and the intermolecular energy.

8.2.7. Valine Mutated Polyalanine Helix

The partial residue energies for the valine mutated polyalanine helix (Hex_Val) system simulation are contained in Table 8.8 and the plots of the values are illustrated across the simulation in Fig. 8.12.

The plots for the partial residue energies indicate that the intra/intermolecular energies are equilibrated, but the intramolecular energies possess larger fluctuations, whereas the intermolecular energy plots possess reduced mobility and oscillation. This result partly agrees with the previous helix structural stability results (Chapter 6) which indicate that the helix is mobile in its core residues and fraying at the ends but the energetic plots do not exhibit a discernible change at about 45 ps, a point in the simulation where residue 29 undergoes a major conformational change to a β conformation.

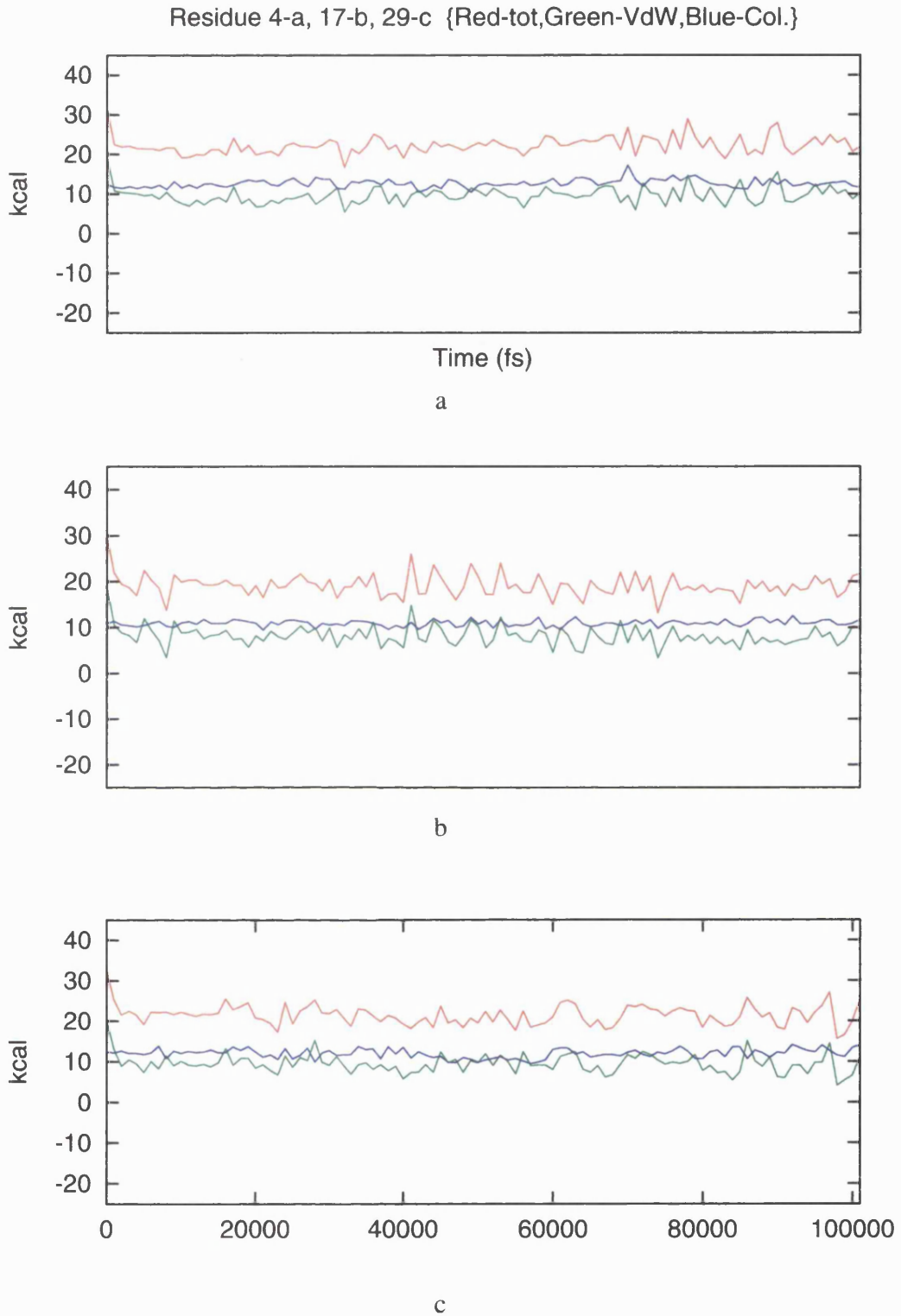


Fig. 8.12 - Hex_Val Intramolecular Partial Residue Energies

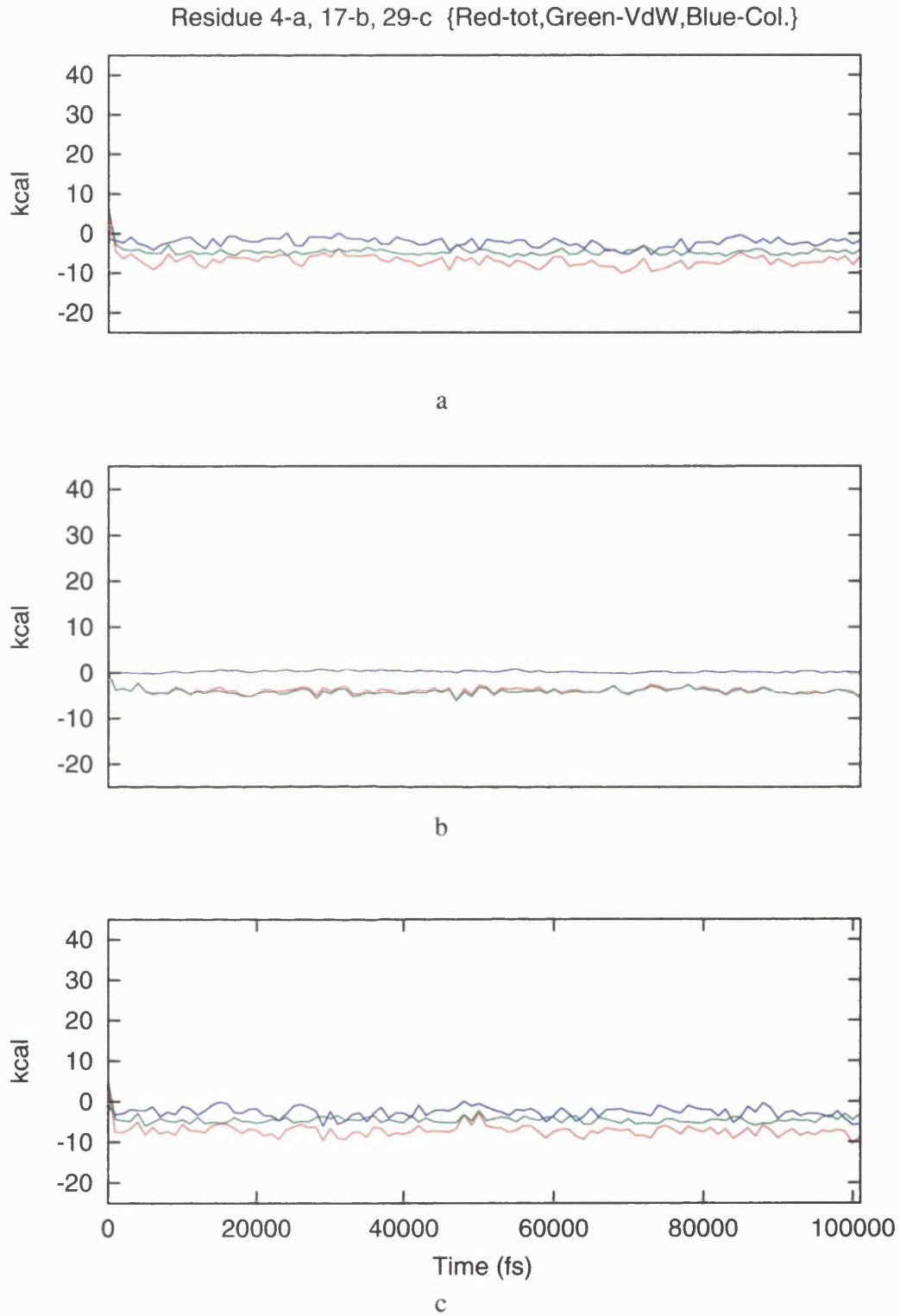


Fig. 8.12 - Hex_Val Intermolecular Partial Residue Energies

This lack of major change either in the intermolecular or more significantly the intramolecular plots indicates that these energies for the valine amino acid are not sensitive to such a change of structure. This is probably a result of the valine residue being aliphatic and having little preference for helix over extended or β structures.

Whilst the plot for the intramolecular energy does not exhibit a feature at 45 ps relating to the conformational change in residue 29, the average values do indicate that there is a difference between the two terminal valines. The surprising result is that the frayed residue 29 has a lower intramolecular energy indicating favours the extended geometry in the interface region of the bilayer. It clearly favours the packing of the β strand when in this environment and is better able to satisfy it hydrogen bonding with the polar solvating atoms than the back bone atoms. This is also an observed feature when comparing the averages for the centrally located residue 17, which has a lower stability than the frayed residue 29. This is interesting as such aliphatic residues are thought to favour the helical geometry in a membrane environment but clearly do not do so in the surface region of the membrane.

Table 8.8

Hex_Val, Helix Partial Residue Energies (kcal mol^{-1})									
	Residue 4			Residue 17			Residue 29		
	total	V.D.W.	coul	total	V.D.W.	coul	total	V.D.W.	coul
Totalmol.	15.42	5.03	10.34	15.29	4.12	11.17	14.27	4.96	9.31
Intramol.	22.27	9.61	12.65	19.12	8.21	10.92	21.43	9.47	11.96
Inter	6.85	-4.58	-2.26	-3.83	-4.09	0.25	-7.16	-4.51	-2.65
				Alanine Residue 5-28					
Totalmol.				5.14	-0.52	5.66			
Intramol.				9.09	2.72	6.37			
Inter				-3.95	-3.25	-0.71			

The intermolecular averages in Table 8.8 again surprisingly indicate that the frayed residue 29 has a more favourable interaction with the surface of the bilayer than the mobile but un-frayed residue 4. As above this result indicates that the valine side

chain favours being solvated in the interface in the frayed conformation over the helical conformation. Thus, the fact that the intramolecular and intermolecular partial residue energies for the frayed, residue 29 are lower than those for the helical residue 4, the valine amino acid favours the β conformation when in the interface of a fully hydrated DMPC bilayer. This result is in keeping with the structural propensity of valine for β-sheet in globular proteins.[1]

The intermolecular energies show that residue 17 has a favourable interaction with the lipid bilayer and there is little difference in the VdW energy in the three positions in the helix. The terminally associated residues have a lower energy than the central mutation site due to the ability of the bilayer interface to hydrogen bond through its polar atoms to the mobile backbone polar atoms. The total energies suggest that in conclusion the frayed β conformer is more stable intermolecularly, than the helical geometry in the interface and in the bilayer core.

Thus, the solvation energy for the α_R helical valine amino acid residue in the hydrophobic core of a full hydrated DMPC bilayer is -3.83 kcal mol^{-1}, with the solvation energy of the α_R helical valine amino acid residue in the interface region is -6.85 kcal mol^{-1} and the solvation energy of a frayed conformer in the interface region is -7.16 kcal mol^{-1}.

The 5-28 average energies for the alanine helix core are contained in Table 8.8 and indicate that the structure of the residues has the essentially the same stability as the core of the pure alanine, Hex_Ala1, helix and this agrees with the previous helix stability results (Chapter 6). Thus, the hydrophobic side chains of the alanine and valine residues are less stabilising at the mutation site, residue 17, than the other more hydrophilic side chains.

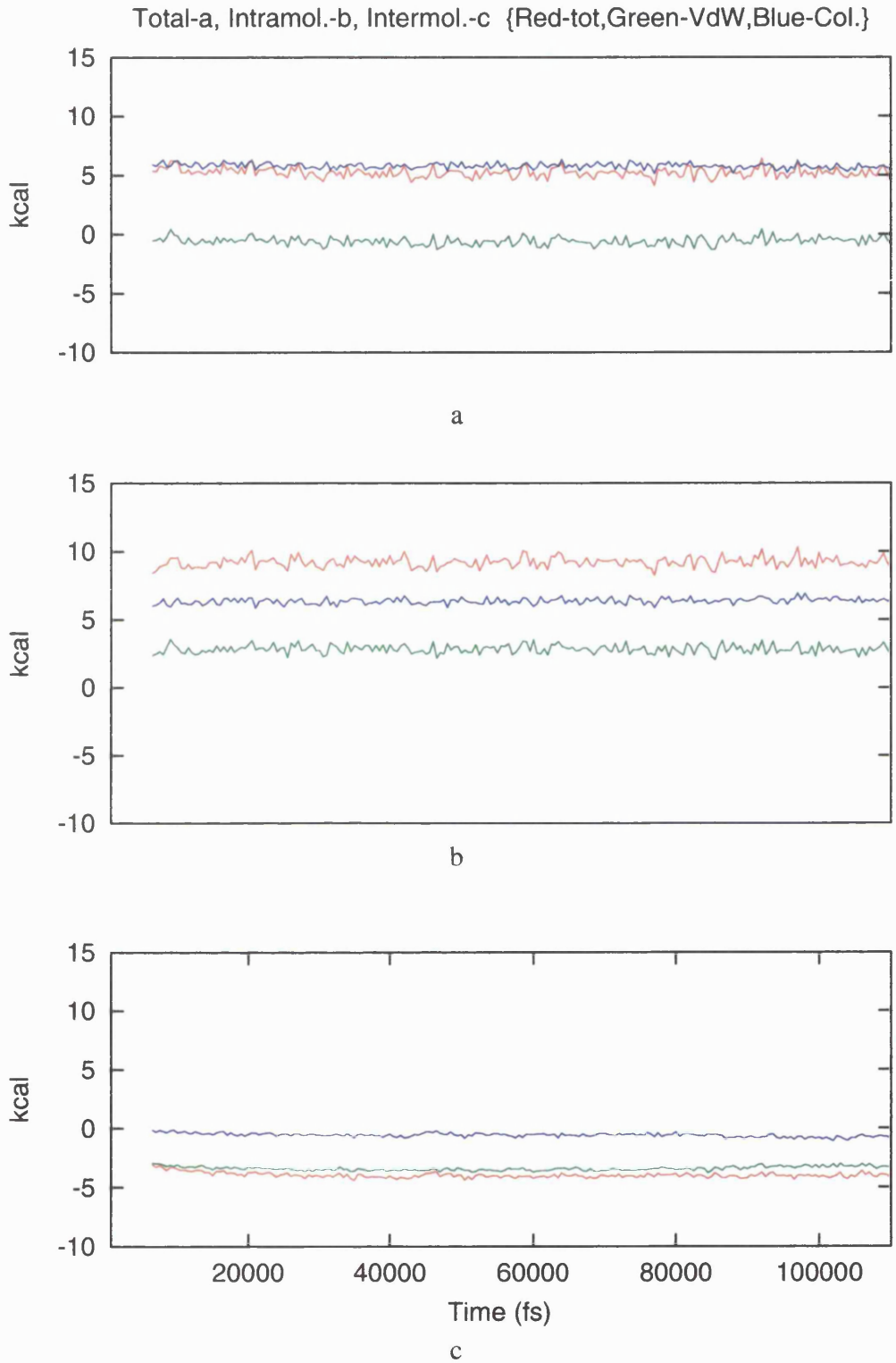


Fig. 8.13 - Hex_Ala2p Helix 1 Partial Residue Energies

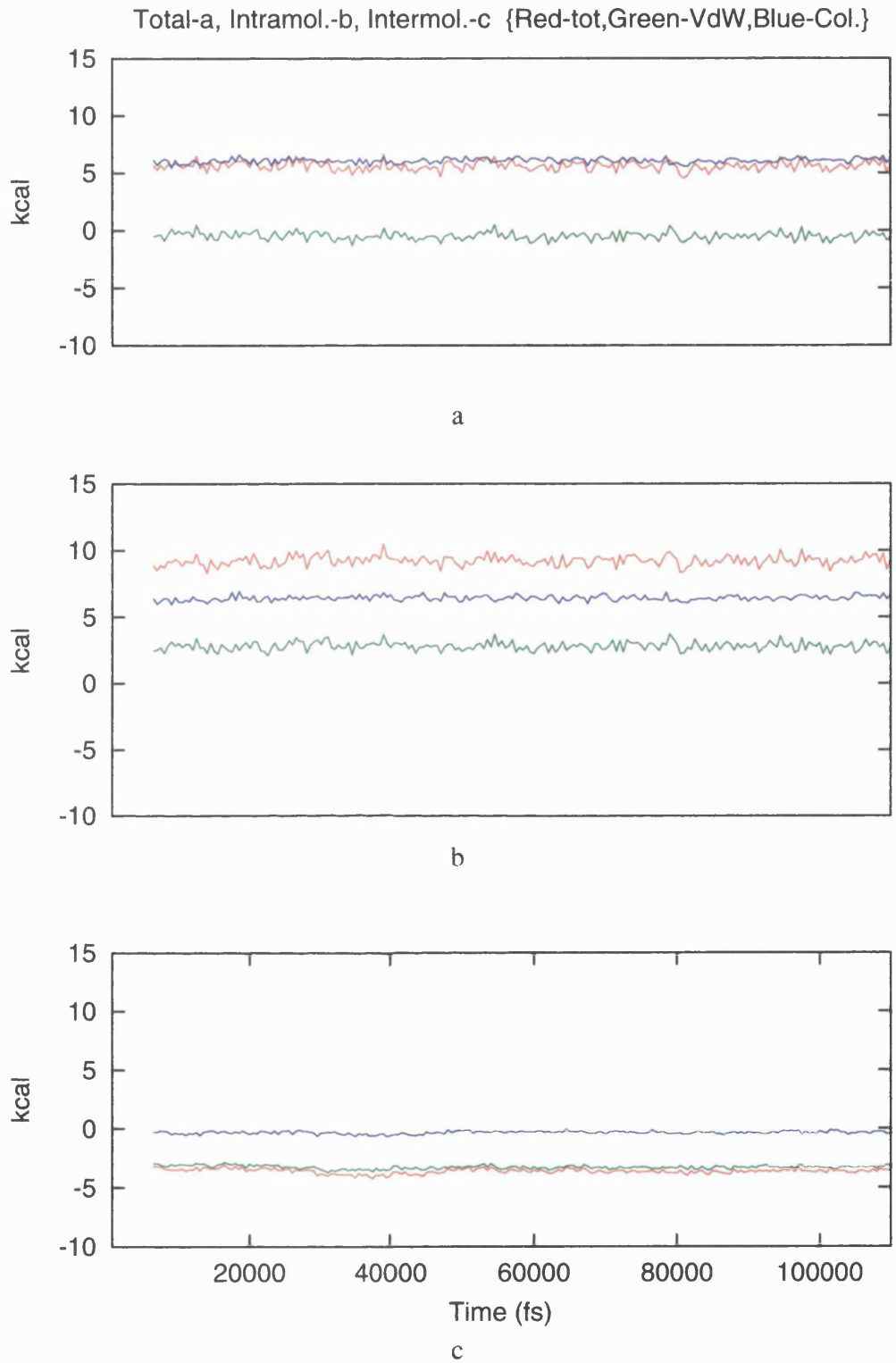


Fig. 8.13 - Hex_Ala2p Helix 2 Partial Residue Energies

8.3. Helix Pair Simulation Results

8.3.1. Parallel Polyalanine Helix Pair

The partial residue energies for the parallel polyalanine helix pair (Hex_Ala2p) system simulation are contained in Table 8.9 and the plots of the values are illustrated across the simulation in Fig. 8.13.

The plots of the energies for the two parallel helices indicate that the intramolecular energy for each helix of the pair is equilibrated and they do not exhibit major changes but fluctuate, as expected for a dynamic system. Conversely the intermolecular energy plots show that changes occur during the simulation in the helix core residue environment, which is in agreement with the helix stability results of Chapter 6. The major observed loss of structure in the Hex_Ala2p helix pair was in the terminal residue, whilst the central residues lost a small amount of helicity. This loss is confirmed to be due the helices optimising their intermolecular interactions. The three energies exhibit changes during the warm up and template forcing section of the simulation, as the plots gradually increase early in the simulation, especially the VDW energy, as a result of the motion of the system.

The observed intermolecular energy optimisation could also be due the lipid bilayer and helices equilibrating their interactions, especially during the initial period of the simulation as the helices become more dynamic. The intermolecular energy of helix 2 clearly becomes more negative between 30 and 45 ps, which corresponds to a loss of helical geometry.

The average intramolecular energy values in Table 8.9 indicates that the core residues of the 2 helices have a similar stability through the 24 residues included in the average. Any differential stability along the helix length appears to average out to give a mean value. Thus, surprisingly the observed dynamic nature of the parallel helices is not replicated in the average intramolecular energies which appear to be

Table 8.9

Hex_Ala2p, Helix Partial Residue Energies - 5-28 Ave. (kcal mol^{-1})						
	Helix 1			Helix 2		
	Total	V.D.W.	coulombic	total	V.D.W.	coulombic
Total	5.22	-0.57	5.79	5.60	-0.48	6.08
Intramol.	9.20	2.83	6.36	9.22	2.81	6.42
Intermol.	-3.98	-3.41	-0.57	-3.62	-3.29	-0.34

resilient to change in local amino acid conformation.

At first glance the intermolecular energy average results (Table 8.9), as mentioned in the observation above, demonstrate differential behaviour between the two helices. The average for helix 1 is lower than that for helix 2 and this appears to suggest that interactions for helix 2 are significantly different than for helix 1, but the intramolecular energy does not indicate structural alterations are occurring in these particular residues. No structural reason was found to cause this but a closer inspection of the bilayer revealed a slight difference in the packing of the lipids around the helices. The packing and lipid density around helix 1 is slightly higher than that around helix 2 which is an artefact of the building of the cavity.

During the building process and the production of the cavity, changes in unit cell vectors (UCV) and thus the volume of the system were used to maintain the density of the lipid system. These changes in the UCV clearly induced a differential density around the two helices which the system was not able to equilibrate sufficiently during the simulation. The inability to remove this density inaccuracy probably results from a combination of the long time scale for such lipid rearrangements and the discontinuity in such lipid motions caused by the inclusion of the helix.

To achieve a more accurate value for the intermolecular energy for residues 5-28 two methods are used to calculate a new average, firstly the two helical values are averaged and secondly, the value for helix 1 is taken as it has a packing density closer to that around the helix in the Hex_Ala1 system. Thus, the first method postulated

above gives a solvation energy for a α_R helical alanine residue in a parallel helix pair, in the hydrophobic core of a fully hydrated DMPC bilayer, -3.80 kcal mol^{-1}. This value is 0.39 kcal mol^{-1} more positive than the average for the same residue in a single helix. The second method gives a solvation energy for helix 1, -3.98 kcal mol^{-1} This is 0.21 kcal mol^{-1} below the value for the same residue in a single helix.

Therefore, taking either value the solvation energy of a alanine helical residue in a parallel helix pair in the bilayer is less favourable than that of the pure alanine single helix. Thus, taking the probably more accurate method 2, the alignment of the helix dipole has a significant effect on each core residue which results in a destabilising effect of 9.36 kcal for the core section of the helix with respect to the pure alanine helix.

8.3.2. Antiparallel Polyalanine Helix Pair

The partial residue energies for the antiparallel polyalanine helix pair (Hex_Ala2a) system simulation are contained in Table 8.10 and the plots of the values are illustrated across the simulation in Fig. 8.14.

Table 8.10

Hex_Ala2a, Helix Partial Residue Energies - 5-28 Ave. (kcal mol^{-1})						
	Helix 1			Helix 2		
	Total	V.D.W.	coulombic	total	V.D.W.	coulombic
Total	5.00	-0.39	5.39	4.91	-0.64	5.55
Intramol.	9.34	2.83	6.51	8.89	2.69	6.20
Intermol.	-4.34	-3.21	-1.12	-3.98	-3.32	-0.65

The plots of the energies for the two anti-parallel helices indicate that the intramolecular energy for the helix pair changes to a more positive value, associated with the removal of the template forcing of the helix geometry at the start of the simulation and the motion of the system. The intramolecular plots, whilst possessing a fluctuating profile, do not exhibit the changes in conformation observed in the helix

stability data (Chapter 5), as was also the case for the Hex_Ala2p energies. Again, this is probably because the changes in the core section structure are located at specific residues and as the other residues do not show significantly altered structure these changes will be smoothed out by the averaging process used to calculated the energy for the residues 5-28.

The intramolecular plot for helix 1 does shows small change in profile that could be associated with a conformational change but in general the intramolecular energy plots are insensitive to structural change. The intermolecular energy plots also show the warming-up of the system and then maintain very stable profiles with no discernible transition in energy. The intermolecular energy for helix 1 starts at a higher value than helix 2 and gradually reduces its energy to a similar equilibration value.

The average intramolecular energy values in Table 8.10 indicate that there is a difference in the interactions of the helices and the lipid bilayer. Helix 1 has a value above that observed for both the Hex_Ala2p helices and the Hex_Ala1 helix and is a result of a loss of helicity in the N-terminal section of the peptide that propagates into the central section, which is not observed to such an extent in the other pure alanine helices. An obvious cause for this major conformational change in the N-terminal and upper section of the helix has not been found and is most likely to occur because of an optimisation of the structure to minimise free space.

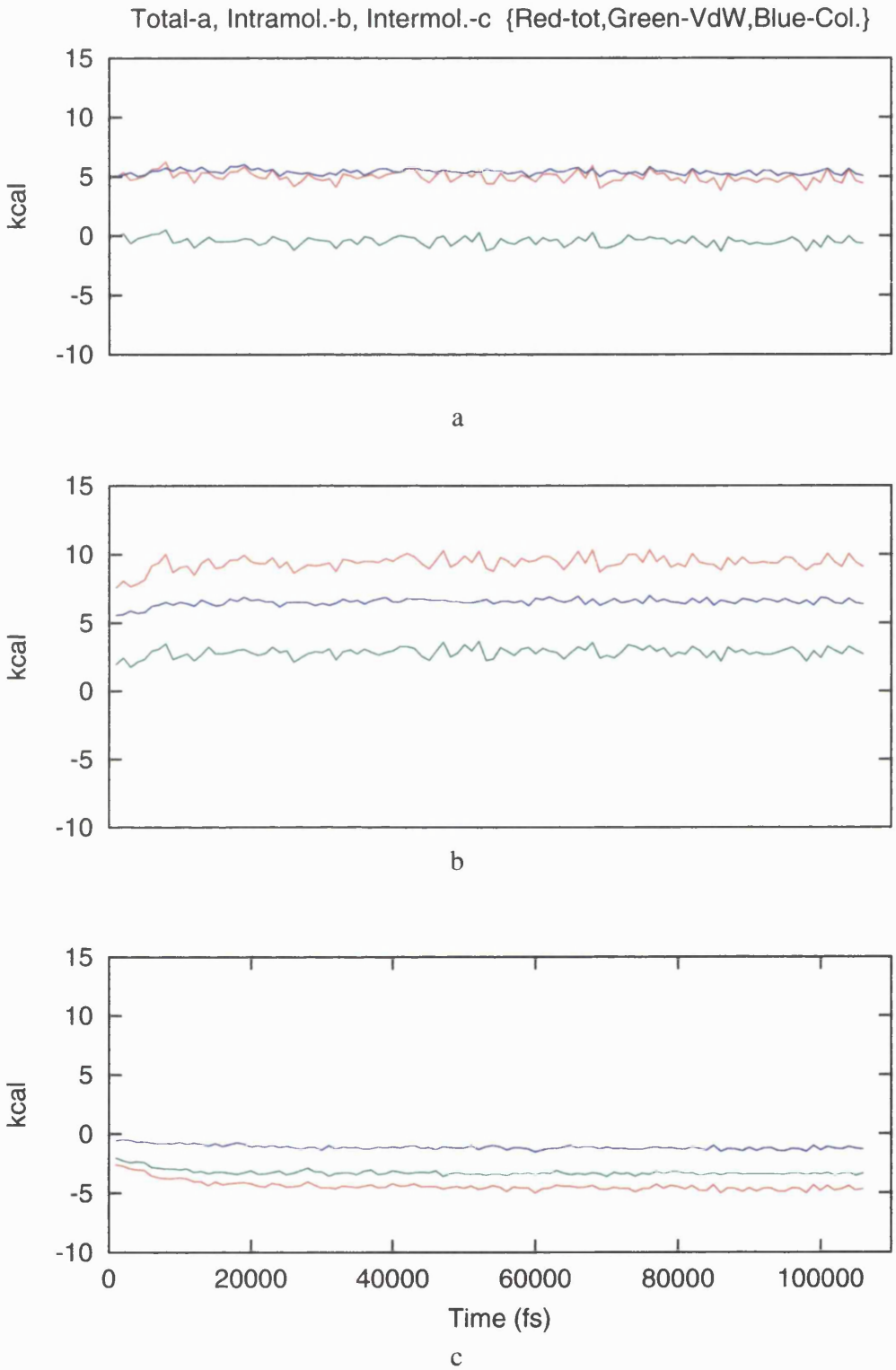


Fig. 8.14 - Hex_Ala2a Helix 1 Partial Residue Energies

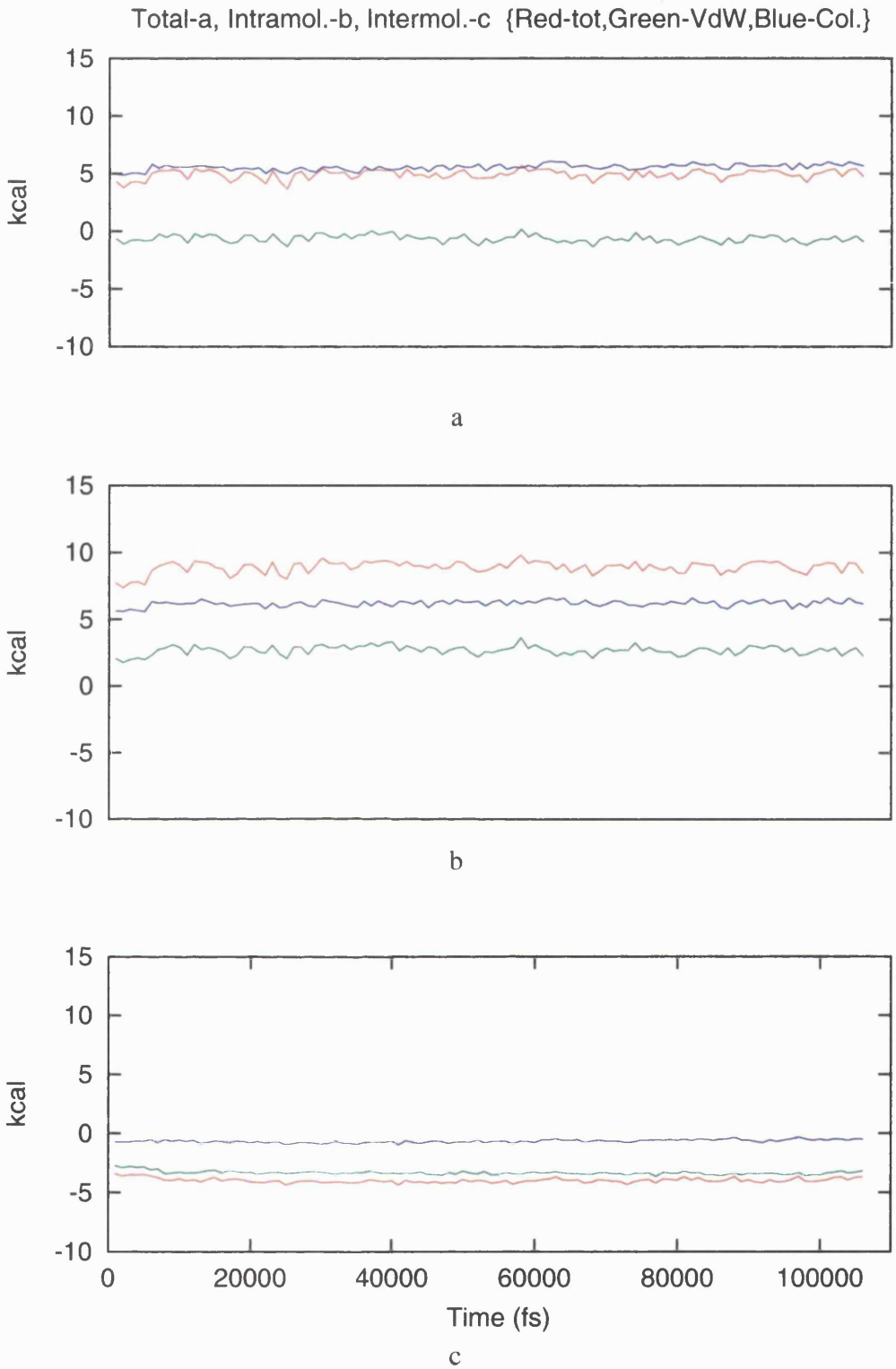


Fig. 8.14 - Hex_Ala2a Helix 2 Partial Residue Energies

The intramolecular energy for helix 2 is typical of the 5-28 averages observed in other stable helices in this study which possess a mutated residue 17 that acts to increase the stability of the core alanine residues. Taking an average of the two helical intramolecular energies gives a value 9.11 kcal mol^{-1} below that of the parallel example and equal to that of the single helix example.

The intermolecular energy averages, as with the Hex_Ala2p simulation, indicate that there is a difference between the environment of the two helices in the pair. Again, the differential in the lipid packing density around the helices is observed. The loss of helical structure in the upper residues in helix 1 results in the backbone satisfying some of its hydrogen bonding requirement intermolecularly and the lower density around the helix also affects the coulombic energy in a favourable direction. The reduced helical geometry in the upper section of helix 1 could also contribute to this differential, as a result of the loss of structure observed early in the simulation.

As with the Hex_Ala2p simulation the differential density has made a clear conclusion difficult to make. As above, two methods of clarifying the results are possible, i.e. firstly by taking a mean of the two helix results and secondly by taking helix 1 as the best example given its lipid packing density is similar to that of the Hex_Ala1 simulation. The method 1 solvation energy for a α_R helical alanine residue in an antiparallel helix pair in the hydrophobic core of a fully hydrated DMPC bilayer is -4.16 kcal mol^{-1} and the method 2 solvation energy is -4.34 kcal mol^{-1}. These values give a change from the single helix values of +0.03 and -0.15 kcal mol^{-1} and a 0.36 kcal mol^{-1} stabilisation over the parallel helix pair. Plus, as the packing around helix 1 is similar to the Hex_Ala1 system then the anti-parallel arrangement can be said to confer 0.15 kcal mol^{-1} of stability over the single helix.

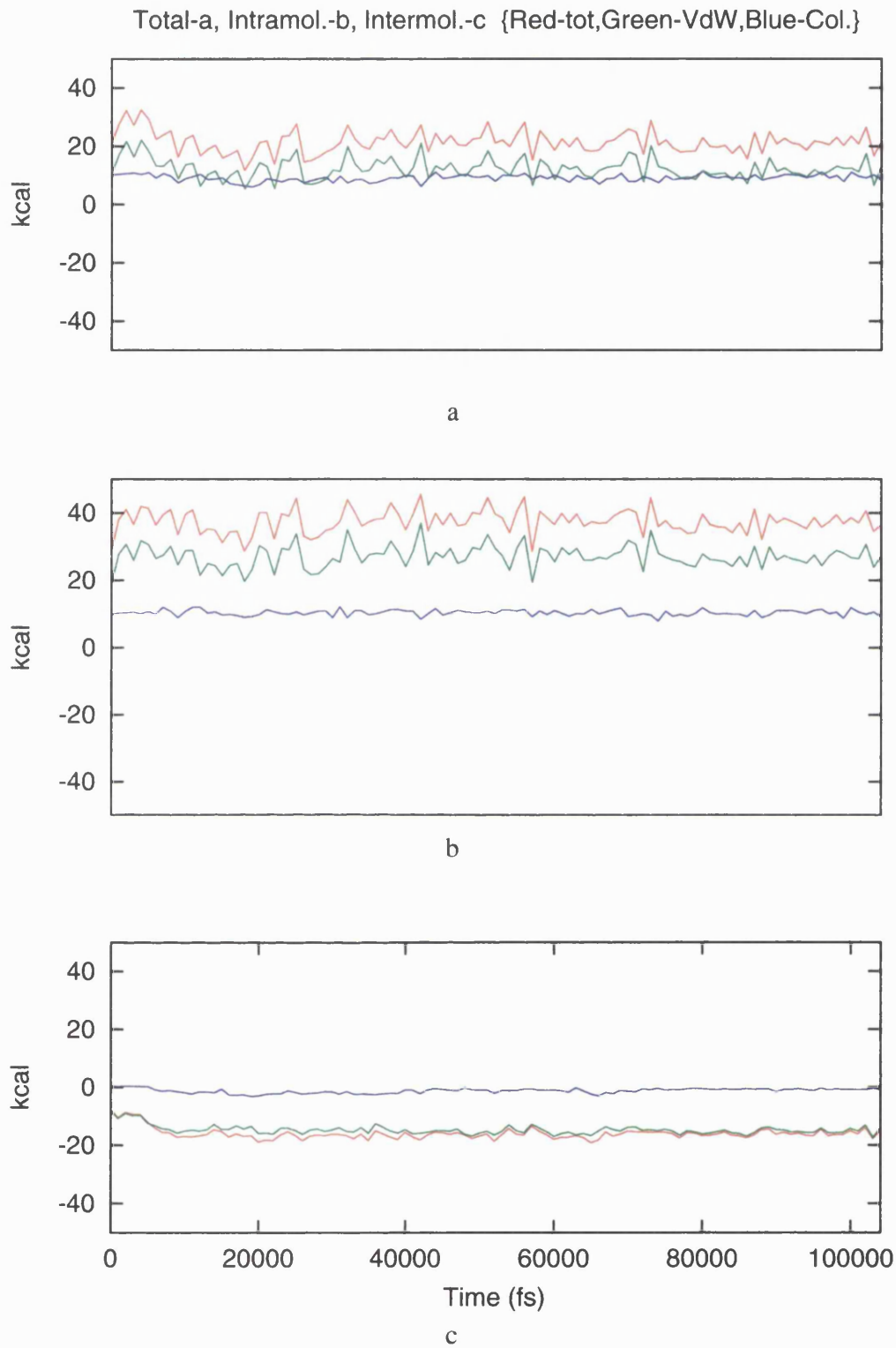


Fig. 8.15 - Hex_Trp_Tyr Helix 1 Partial Residue Energies

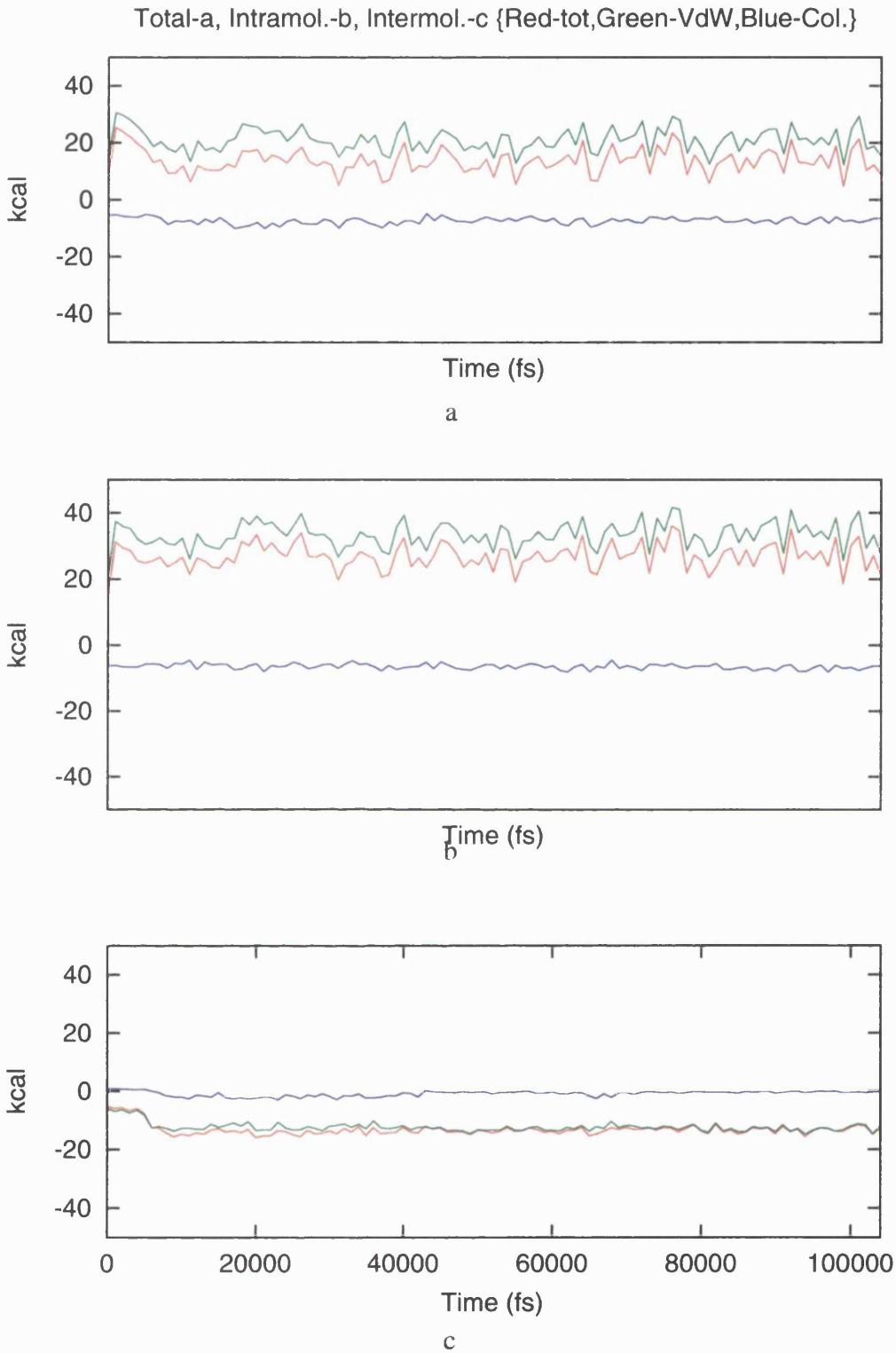


Fig. 8.15 - Hex_Trp_Tyr Helix 2 Partial Residue Energies

8.3.3. Tryptophan and Tyrosine Mutated Antiparallel Polyalanine Helix Pair

The partial residue energies for the tryptophan and tyrosine mutated antiparallel polyalanine helix Pair (Hex_Trp_Tyr) system simulation are contained in Table 8.11 and the plots of the values are illustrated across the simulation in Fig. 8.15.

The plots of the energies for the two anti-parallel helices indicate that the intramolecular energy for the two helices have equilibrated from 20 ps onwards, after a period where the helices appear to be minimising their structure. The tryptophan residue loses a significant amount of VdW energy as it arranges its structure, most probably as it adjusts the geometry of the side chain as the backbone data indicates no such change. This is also true for the tyrosine of helix 2, where the intramolecular energy exhibits a dip in the profile at 20 ps, from then on it fluctuates around its equilibrium value.

The intermolecular energy plots indicate that the conformational changes in the amino acid side chains have significant effects on their interactions. Clearly the rearrangement of the structure was not just to minimise the geometry of the side chains but more pertinently to minimise the packing of the two ring systems. This geometry significantly stabilises the residues and the pi-rings become more aligned, with the subsequent slight change in the electrostatic energy associated with stacked ring systems.

The intermolecular residue partial energy values in Table 8.11 indicate that both the tryptophan and the tyrosine residues have favourable enthalpies in the anti-parallel helix pair in a bilayer environment. Thus, the formation of such partially aligned aromatic amino acids will favour the aggregation of helices into bundles over the related single helix example. The favourable nature of the Phe amino acid in the single helix simulation, which gave a bilayer solvation energy of -8.61 kcal mol^{-1}, contrasts with that for the two related tryptophan and tyrosine residues which have stabilising energies of -16.03 kcal mol^{-1} and -13.28 kcal mol^{-1}. Thus, the pairing of the two aromatic

Table 8.11

Hex_Trp_Tyr, Helix Partial Residue Energies (kcal mol^{-1})						
	Helix 1 (Trp)			Helix 2 (Tyr)		
	total	V.D.W.	coulombic	total	V.D.W.	coulombic
Total	21.01	11.90	9.11	13.34	20.60	-7.26
Intramol.	37.04	26.66	10.38	26.62	32.97	-6.35
Intermol.	-16.03	-14.76	-1.27-13.28	-12.37	-0.91	
	Alanine Residue 5-28					
	Helix 1			Helix 2		
Total	5.20	-0.49	5.68	5.29	-0.51	5.80
Intramol.	9.09	2.72	6.37	9.01	2.66	6.35
Intermol.	-3.90	-3.21	-0.68	-3.72	-3.17	-0.55

residues, even though they are not identical but are very closely related, gives the amino acid between 4.67 and 7.42 kcal mol^{-1} of additional stabilisation.

This stabilisation of paired aromatic residues means that the phenylalanine residue's preference for the interface, previously observed, could be altered in favour of pairing with a phenylalanine from another helix in the bilayer interior. This is because the intermolecular energy for residue 4 and residue 29 in the Hex_Phe simulation are 5.21 and 6.21 kcal mol^{-1} more negative than residue 17 but taking into account the observed aromatic pairing stabilisation effect of between 4.67 and 7.42 kcal mol^{-1} it is possible that this type of aromatic coupling will cause a stabilisation of the paired phenylalanines over an isolated interface associated phenylalanine.

To better improve the evaluation of the aromatic pair stabilisation energy this simulation should be repeated using a pair of Hex_Phe helices and not the Hex_Trp_Tyr pair used here in order to investigate the observed aromatic pair from Bacteriorhodopsin.

The 5-28 alanine intramolecular average energies in Table 8.11 indicate that the structure of the core alanine residues are similar to that of the pure alanine simulation ,Hex_Ala1, but the intermolecular energy indicates that the lipid and inter-helix

interactions are less favourable, more like those from the parallel helices in Hex_Ala2p. The energy reflects the structural changes induced in helix 1 by the helices reducing their separation and the rearrangement of the two ring systems. Thus, whilst the anti-parallel arrangement of the helices previously had a stabilising effect on the helices, the rearrangements induced by the aromatic residues disrupts the inter-helical interactions.

This effect is also a consequence of the build induced artefact in the system density which causes the packing of the lipids to be lower around helix 2. The major destabilisation of the helix structures is due to the need to separate the helices in the anti-parallel pair in order to accommodate the mutation from alanine to tryptophan and tyrosine.

8.4. Discussion

The results above highlight several interesting properties of helices in lipid bilayer environments and give energies that can be used to explain structural behaviour and physically observed behaviour of membrane peptides. This will aid in the energetic description of protein structural properties in membrane environments.

Table 8.12

Summary of E_{int} Results		
System	H.Gp.	Ac.
Hex_Ala1		-4.19
Hex_Thr	-11.87	-3.69
Hex_Val	-7.00	-3.83
Hex_Asp-	-80.66	0.99
Hex_Asp	-16.29	-5.00
Hex_Arg	-61.73	-38.72
Hex_Phe	-14.82	-8.61
Hex_Tyr		-12.37
Hex_Trp		-16.03
ΔE_{aro}		-6.04
Hex_Ala2p		-3.98
Hex_Ala2a		-4.34
ΔE_{par}		-0.36

E_{int} is the interaction energy of the residue
ΔE_{par} is the parallel versus anti-parallel enthalpy
ΔE_{aro} is the aromatic stabilisation energy
Ac. - Residue 17, H.Gp. - Average of Residues 4 and 29
All energies in kcal mol^{-1}

Interestingly the partial residues energetics indicated that the intramolecular energies are insensitive to modifications in the internals of one residue and any energetic change that results from such a modification must be compensated for or swamped by changes in other residues so as to maintain the average energy. The changes in the intramolecular energy averages for the single helix simulations indicates that relatively large changes in structure only produce small changes in the averages.

The mutation of the central residue to the other examples used here generally stabilises the helix core region, except in the Hex_Asp- system which is also more soluble in the bilayer. Other than this slight increase for a charged Asp, the core alanine residues favour the incorporation of other residues as they increase the structural stability of the helix at a cost of reduced interactions with the lipid molecules. The most structurally stabilising mutation is Arg, with an increase in core stability energy of 0.52 kcal mol^{-1} and a reduction in solvation energy of 0.31 kcal mol^{-1}.

Interestingly the more polar or charged residues have the most affect on the helix structure as they increase helix stability rather than decrease it as might be expected. Thus, the inclusion of polar or charged entities into a bilayer incorporated, α_R, helix results in a more favourable electrostatic energy which is a consequence of the strengthening of the back bone hydrogen bonding interactions that holds a membrane bound helix together and the factor that makes the helix a predominant secondary structure. Interestingly they have the effect of increasing the solvation energy as they maximise their internal interactions.

Also, the most hydrophobic residue valine, with its branched chain aliphatic structure, has the least structurally stabilising effect on the core alanine residues of all the mutated helices. It also reduces the solvation enthalpy of the helix core when compared to the Hex_Ala1 simulation. This is surprising given that it is accepted that helices containing hydrophobic residue are designated as more likely to incorporate into a bilayer environment. Even given the favourable solvation energy of the valine residue in the interior of the bilayer the most hydrophobic helix does not have a significantly lower solvation enthalpy than the polar helices, Hex_Val being essentially the same as Hex_Thr.

The effect of charged residues on the energetics and stability of a helix in a bilayer environment is a central topic of interest to membrane protein science. A major question is whether such residues are actually charged in such a non-polar

environment?

The charged aspartate residue has a small positive solvation energy, thus has little or no preference for inclusion in the bilayer. When compared to the large favourable negative energy of the residues in the interface the inclusion of such a charged entity in the bilayer is unlikely and movement form the membrane surface to the interior in not preferred. The neutral aspartic acid side chain has a much more favourable solvation enthalpy in the bilayer interior but again the most favourable position is in the interface region. Thus, an aspartate residue whether charged or neutral has a far higher affinity for the surface region than the interior of a membrane and will most likely exist in the de-protonated form in the bilayer interior.

The other single helix simulations all give a favourable solvation energy for the residue 17 mutation site as they all exhibit negative partial residue intermolecular energies. What is interesting is that the arginine residue which would be expected to disfavour inclusion, as it is as strongly charged as the aspartate residue, is able to drastically improve its solvation enthalpy by the rearrangement of its side chain so its hydrophilic atoms are directed towards the head group associated water. Further, the side chain is then capable of attracting head group waters into the hydrophobic part of the bilayer. Hence the long and flexible nature of the arginine side chain means that even buried deep in the bilayer, it can achieve a dramatic stabilisation due to fluid nature of the lipid bilayer surface.

The anti-parallel helix pair causes the core alanine residues to be 0.15 kcal mol^{-1} more stable than in the single helix pure alanine and the parallel helix is 0.21 kcal mol^{-1} less stable. Thus, the anti-parallel helix arrangement is shown to be more favourable as is expected.

The second possible factor for helix pair stabilisation studied is the formation of an aromatic pair, where the alignment of the rings causes the stabilisation of the system. The example investigated here indicates that this stabilisation is quite significant

and could be in the range of between 4.67 and 7.42 kcal mol^{-1}. Of this stabilisation a small proportion is due to the anti-parallel arrangement of the helices, but this reduction still leaves a large degree of stabilisation resulting from the pairing of the aromatic residues.

On comparing these results with those of Woolf et al[2] it can be seen that our results are similar for certain combinations of residues and positions and less so for others. His results (Page 19 - Introduction) combine a number of different positions around the center or interface, depending on the random placing of amino-acids in the bacteriorhodopsin sequence rather than the specific position reported here. Further, the use of the bacteriorhodopsin sequence means the energies for each amino-acid type are defined in the field of the other amino-acids, rather than the constant background of Alanine used here.

Thus, these detailed results give new and interesting insights into the factors affecting bilayer incorporated helix behaviour. The calculated energies begin to give quantitative results for these often observed helix properties and also, for new properties such as the interface water stabilised, centrally located arginine residue. This energetic evaluation of the mutated examples of the alanine α_R helix can now form the first step to the formulation of a potential energy force field that can describe the behaviour of amino acids in the bilayer environment and to thus probe the structural tendencies of membrane protein primary sequences.

References

1. G. D. Fasman, *Prediction of Protein Structure,* Plenum Press, New York (1989).

2. T.B. Woolf, *Biophysical Journal*, **74**, 115-131 (1998).

9. Conclusion

Atomic level detail on the energetics of specific amino acid residues interactions with a hydrated bilayer have been obtained. These solvation energies give a clear insight into the interactions that stabilise helical peptides and paired helices in the mixed dielectric environment of the lipid bilayer. Prior to this bilayer simulations were used to determine the best way to build and equilibrate model bilayers as this is crucial to the validity of our further simulations of the bilayer plus additional molecules.

A preliminary study was carried out to investigate the ability of the VFF to replicate the experimentally observed L_{α}, liquid crystal, phase for the hydrated DMPC bilayer. It was shown that the hexagonal geometry gave the best model membrane of the lipid environment due to its off-set lattice arrangement, whereas the cubic lattice arrangement imparted serious artefacts into the bilayer structure. The hexagonal geometry reproduced all the desired bilayer properties except the correct gauche population of the two acyl chains. This is due to both the refinement procedure, which caused the acyl chains to achieve fluidisation through a mixture of isomerisation and acyl chain scissoring, and the slightly high torsional energy barrier for the hydrocarbons, which reduces the rate of isomerisation from that required. Thus, in future a reparameterisation and longer time simulations will be required to allow the correct gauche population to develop. These factors are independent of the care taken during construction and equilibration.

The incorporated helices maintained a good α helical geometry, with only the terminal residues affected by increased dynamic motion and fraying as observed experimentally. The main structural destabilisation was due to the charged aspartate residue, whilst the majority of the amino acids mutations conferred structural stability on the helices.

The bilayer itself reacted in interesting ways to the helices incorporated into it. Firstly, the single helices reduced the fluidity of the acyl chain via reduced Sn2 isomerisation and scissoring, whilst the helix pair imparted similar Sn2 and scissoring motional dampening but also reduced the isomerisation of the Sn1 acyl chain. Thus, the helices acted to reduce the conformational space of the acyl chains.

The energetic analysis is the main focus of this study and is a precursor to the development of solvation parameters for the bilayer environment. The next phase of this process will be to study the remaining amino acids in order to get a full set of lipid solvation energies. With the ever increasing computer power this task will be much easier. Then the same host/guest peptides will be run through aqueous phase simulations in order to calculate hydration energies, which will permit the calculation of the partition enthalpies. The culmination of this work will be the development of a mean force field, with lipid solvation parameters, that will allow the folding of membrane proteins from their amino acid sequence into a folded tertiary structure with the explicit lipid solvation parameters.

Obviously, this stage of the work is sometime from being realised and thus, we have been working on a new intermediate methodology that combines the relatively small amount of experimental data and structural detail available for a membrane protein with model building techniques develop in house. The method is outlined in Appendix I (paper in preparation) and has been used to build a model for the G-Protein Coupled Receptor super, 5-HT1_A. The receptor model produced by the semi-automated model production scheme is presently being studied using conformational sampling techniques and MD in order to refine the structure with a view to membrane inclusion and docking studies (Fig. 9.1 shows the 5-HT1_A model with cleft hydration, using explicit waters, and docked serotonin).

It is hopes that as the work progresses these two methods can be refined and merged so as to eventually produce a combined empirical and a-priori folding

methodology that will fill the gap between the dearth of structural detail and requirements of membrane protein science.

Fig. 9.1 - 5-HT1$_A$ Receptor Model with Cleft Waters and Docked Serotonin

Appendix I.

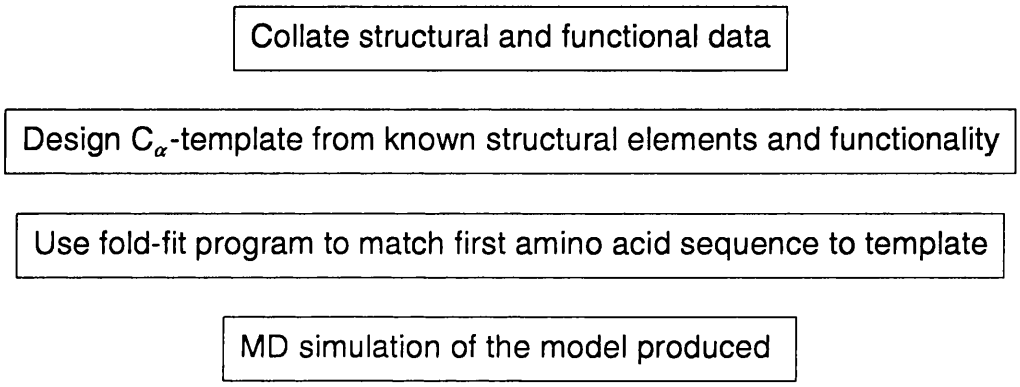
G.P.C.R. Model Production Technique

As discussed in Chapter 1 the difficulties of acquiring experimental data on membrane protein structure has focused attention on new methodologies to side-step this problem. We have developed a new feature of the V.F.F. programme which enables a protein or peptide structure to be built using the available secondary structure and mechanistic information. This method has been used to produce a new model for the G-protein coupled receptor, $5-HT_{1A}$. The basis for the development of this process is discussed below and then the process that is carried out to produce a 3D structure for the receptor will be outlined.

The modelling process (Fold Fit) involves using the bank of information on bioaminergic receptors, such as theoretical considerations of the likely protein folding patterns within a membrane, the large body of experimental data on the topological orientation of the helices and the ligand-binding properties for these receptors. The method uses a combination of molecular dynamics simulations and energy minimisations, along with atomic coordinate and dihedral angle constraints, to map the full receptor amino acid sequence onto a pre-modelled C_α atom template of the putative seven transmembrane helix regions. This methodology allows amino-acid sequences to be added section by section with a pre-determined secondary structure for the known or predicted structure of the protein, with the inter-helical loops, side chains and terminal domains being built without predetermined structure. It is hoped that such models will act as sensible structures for the analysis of the dynamic properties of these proteins.

Fold Fit Methodology

The semi-automated Fold Fit procedure is outlined in the schematic below and the individual steps are briefly discussed. The literature information used to build the template model is highlighted for the $5-HT_{1A}$ example.



Schematic of the Step-wise Fold Fit Procedure

Collation of Data

Sequence Alignments:

Sequence alignment of the seven regions for the $5-HT_{1A}$ receptor using PHD along with sequence portions of the TMs I-VII published elsewhere were used.[1-3]

Conserved Residues:

The rhodopsin-like receptor super family, including $5-HT_{1A}$, is characterised by the conserved residues: GN in TM1, LXXXD in TM2, DRY at the end of TM3 and NP in TM4.[4]

Trans-membrane Structure:

It is known that the binding site of neurotransmitter receptors is formed among their seven mostly hydrophobic segments [5] and the projection density map of rhodopsin shows seven peaks of density that are thought to represent seven TM

helices.[6]

TM Orientation:

The seven peaks of density mentioned above are thought to represent four trans-membrane helices orientated nearly perpendicular to the membrane with the elongated arc-shaped features probably representing the three remaining in tilted positions.[6]

Hydrophilic Residue Orientation:

The TM segments include mainly hydrophilic residues or residues with polar side chains and probably will be found in the region considered as extra-membranous and have been used as an indication of the possible limits of TM sections. The occurrence of hydrophilic residues in the TM section was used as an indicator of those parts that will probably be orientated into the protein interior.

TM Section Length:

The length of the helices spanning the membrane was restricted to the published data value of 26 amino-acids. [6]

Cysteine Bridge:

A cysteine bridge has been reported for $5-HT_{1A}$ and exists between conserved residue 110 at the top of helix 3 and conserved residue 188 on the third extracellular loop.

C_α Template Construction

The information above was used to build a C_α atomic template for the predicted 7TM helices, with a manual fit of the C_α atoms to best replicate the collated information, maximising the proposed stabilising effects and optimising the arrangement of the putative receptor associated residues.

Fold Fit V.F.F. Utility

The new utility of the V.F.F. allows the folding of a primary amino sequence on to a pre-defined C_α template for known or predicted structure using a combination of minisation and MD.

The structure is built step by step from the N-terminus to the C-terminus in predetermined sections of sequence or, for important sites in the protein, via individual residues. The Fold Fit utility builds the sequence in predetermined chunks (stage 1) and then translates the receptor C_α atoms onto their template equivalent C_α atoms (stage 2).

Stage 1

During the building of the structure the pre-existing sections have their template equivalent C_α atoms restrained to the template coordinate and their dihedral angles restrained to prevent the already built loops and N-terminus clashing with the newly added structure. A control file instructs the V.F.F. which residues are to be built at each iteration of the procedure. In order to speed up the process the loops and terminal sections are built as a block from the last residue of a helix to the first residue of the next helix. The first four residues of a helix are built individually so as to ensure the helix is directed correctly down the axis of the equivalent helix in the template. The structure is built with set secondary structure so as to minimise the number of badly built residues, with the TM section set as α_R helices and the non-TM section as

β-strand.

Stage 2

The part built structure is then run through a restraining MD simulation where the C_α atoms of the TM section are forced to the coordinates of the C_α template, with the non-TM sections and all the amino acid side chains being unrestrained. The TM helices also have their backbone torsions restrained to the built values so as to prevent the formation of distorted omega dihedral angles as the C_α atoms are 'dragged' to their new positions. The template forcing of the C_α atoms was done in the MD methodology, as it tends to find a more global optimum than the minimiser. The process was repeated until the C-terminal section was built.

Appendix II

Atomic Charges and Potential Types

NB - The potential library used during the course of this series of simulations is the published version of the VFF[7]

Alanine		
Atom	Potential	Charge
N	n	-0.50
HN	hn	0.28
CA	ca	0.12
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c3	-0.30
HB1	h	0.10
HB2	h	0.10
HB3	h	0.10

Aspartic Acid		
Atom	Potential	Charge
N	n	-0.50
CA	ca	0.12
HN	hn	0.28
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c2	-0.20
HB1	h	0.10
HB2	h	0.10
CG	c'	0.38
OD1	o'	-0.35
OD2	oh	-0.38
HD2	ho	0.35

Aspartate		
Atom	Potential	Charge
N	n	-0.50
CA	ca	0.12
HN	hn	0.26
HA	h	0.07
C	c'	0.38
O	o'	-0.41
CB	c2	-0.26
HB1	h	0.07
HB2	h	0.07
CG	c-	0.34
OD1	o-	-0.57
OD2	o-	-0.57

Pheylalanine		
Atom	Potential	Charge
N	n	-0.50
HN	hn	0.28
CA	ca	0.12
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c2	-0.20
HB1	h	0.10
HB2	h	0.10
CG	cp	0.00
CD1	cp	-0.10
HD1	h	0.10
CE1	cp	-0.10
HE1	h	0.10
CZ	cp	-0.10
HZ	h	0.10
CE2	cp	-0.10
HE2	h	0.10
CD2	cp	-0.10
HD2	h	0.10

Arginine		
Atom	Potential	Charge
N	n	-0.50
HN	hn	0.28
CA	ca	0.12
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c2	-0.20
HB1	h	0.11
HB2	h	0.11
CG	c2	-0.20
HG1	h	0.13
HG2	h	0.13
CD	c2	-0.09
HD1	h	0.13
HD2	h	0.13
NE	n1	-0.50
HE	hn	0.36
CZ	cr	0.45
NH1	n2	-0.50
HH11	hn	0.36
HH12	hn	0.36
NH2	n2	-0.50
HH21	hn	0.36
HH22	hn	0.36

Valine		
Atom	Potential	Charge
N	n	-0.50
CA	ca	0.12
HN	hn	0.28
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c1	-0.10
HB	h	0.10
CG1	c3	-0.30
HG11	h	0.10
HG12	h	0.10
HG13	h	0.10
CG2	c3	-0.30
HG21	h	0.10
HG22	h	0.10
HG23	h	0.10

Threonine		
Atom	Potential	Charge
N	n	-0.50
CA	ca	0.12
HN	hn	0.28
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c1	-0.07
HB	h	0.10
OG1	oh	-0.38
HG1	ho	0.35
CG2	c3	-0.30
HG21	h	0.10
HG22	h	0.10
HG23	h	0.10

Tryptophan		
Atom	Potential	Charge
N	n	-0.50
CA	ca	0.12
HN	hn	0.28
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c2	-0.20
HB1	h	0.10
HB2	h	0.10
CG	c5	0.00
CD1	c5	0.01
NE1	np	-0.50
CE2	c5	0.11
CD2	c5	0.00
HD1	h	0.10
HE1	hn	0.28
CE3	cp	-0.10
HE3	h	0.10
CZ3	cp	-0.10
HZ3	h	0.10
CH2	cp	-0.10
HH2	h	0.10
CZ2	cp	-0.10
HZ2	h	0.10

Tyrosine		
Atom	Potential	Charge
N	n	-0.50
CA	ca	0.12
HN	hn	0.28
HA	h	0.10
C	c'	0.38
O	o'	-0.38
CB	c2	-0.20
HB1	h	0.10
HB2	h	0.10
CG	cp	0.00
CD1	cp	-0.10
HD1	h	0.10
CE1	cp	-0.10
HE1	h	0.10
CZ	cp	0.03
OH	oh	-0.38
HH	ho	0.35
CE2	cp	-0.10
HE2	h	0.10
CD2	cp	-0.10
HD2	h	0.10

Water		
Atom	Potential	Charge
O1	o*	-0.8200
H11	h*	0.4100
H12	h*	0.4100

DMPC (Part 1)					
Atom	Potential	Charge	Atom	Potential	Charge
C	cn	-0.0800	O4	o	-0.3800
HC1	h	0.1300	C8	c'	0.3800
HC2	h	0.1300	O5	o'	-0.3800
HC3	h	0.1300	C9	c2	-0.2000
N	n3	-0.1600	H91	h	0.1000
C1	cn	-0.0800	H92	h	0.1000
H11	h	0.1300	C10	c2	-0.2000
H12	h	0.1300	H101	h	0.1000
H13	h	0.1300	H102	h	0.1000
C2	cn	-0.0800	C11	c2	-0.2000
H21	h	0.1300	H111	h	0.1000
H22	h	0.1300	H112	h	0.1000
H23	h	0.1300	C12	c2	-0.2000
C3	cn	-0.0300	H121	h	0.1000
H31	h	0.1300	H122	h	0.1000
H32	h	0.1300	C13	c2	-0.2000
C4	c2	-0.0400	H131	h	0.1000
H41	h	0.0600	H132	h	0.1000
H42	h	0.0600	C14	c2	-0.2000
O	o	-0.4800	H141	h	0.1000
P	p	1.3000	H142	h	0.1000
O1	o'	-0.7500	C15	c2	-0.2000
O2	o'	-0.7500	H151	h	0.1000
O3	o	-0.4800	H152	h	0.1000
C5	c2	-0.0400	C16	c2	-0.2000
H51	h	0.0600	H161	h	0.1000
H52	h	0.0600	H162	h	0.1000
C6	c1	0.1800	C17	c2	-0.2000
H6	h	0.2000	H171	h	0.1000
C7	c2	0.1800	H172	h	0.1000
			H71	h	0.1000
			H72	h	0.1000

DMPC (Part 2)					
Atom	Potential	Charge	Atom	Potential	Charge
C18	c2	-0.2000	C27	c2	-0.2000
H181	h	0.1000	H271	h	0.1000
H182	h	0.1000	H272	h	0.1000
C19	c2	-0.2000	C28	c2	-0.2000
H191	h	0.1000	H281	h	0.1000
H192	h	0.1000	H282	h	0.1000
C20	c2	-0.2000	C29	c2	-0.2000
H201	h	0.1000	H291	h	0.1000
H202	h	0.1000	H292	h	0.1000
C21	c3	-0.3000	C30	c2	-0.2000
H211	h	0.1000	H301	h	0.1000
H212	h	0.1000	H302	h	0.1000
H213	h	0.1000	C31	c2	-0.2000
O6	o	-0.3800	H311	h	0.1000
C22	c'	0.3800	H312	h	0.1000
O7	o'	-0.3800	C32	c2	-0.2000
C23	c2	-0.2000	H321	h	0.1000
H231	h	0.1000	H322	h	0.1000
H232	h	0.1000	C33	c2	-0.2000
C24	c2	-0.2000	H331	h	0.1000
H241	h	0.1000	H332	h	0.1000
H242	h	0.1000	C34	c2	-0.2000
C25	c2	-0.2000	H341	h	0.1000
H251	h	0.1000	H342	h	0.1000
H252	h	0.1000	C35	c3	-0.3000
C26	c2	-0.2000	H353	h	0.1000
H261	h	0.1000	H351	h	0.1000
H262	h	0.1000	H352	h	0.1000

References

1. F.O. Levy, T. Guderman, M. Birnbaumer, A.J. Kaumann, and L. Birnbaumer, *FEBS Lett.*, **296**, 201-206 (1992).

2. S. Trumpp-Kallmeyer, J. Hoflack, A. Bruinvels, and M. Hibert, *J. Med. Chem.*, **35**, 3448-3462 (1992).

3. W. Kuipers, H. J. Kruse, I. van Wijngaarden, P.J. Standaar, M.T.M. Tulp, N. Veldman, A. K. Spek, and A. Ijzerman, *J. Med. Chem.*, **40**, 300-312 (1997).

4. L. Oliveira, A.C.M. Paiva, and G. Vriend, *J. Comp. Mol. Desgn.*, **7**, 649-658 (1993).

5. D. Donnelly and J.B.C. Findlay, *Current Opinion in Structural Biology*, **4**, 582-589 (1994).

6. J.M. Baldwin, *EMBO J.*, **12**, 1963-1703 (1993).

7. P. Dauber-Osguthorpe, V. Roberts, D. J. Osguthorpe, J. Wolff, M. Genest, and A. T. Hagler, *Protein : Structure Function and Genetics*, **4**, 31-47 (1988).